

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



## **Molecular Genetic Analysis of Hidradenitis Suppurativa (Acne Inversa)**

Pink, Andrew Edward

*Awarding institution:*  
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

### **END USER LICENCE AGREEMENT**



**Unless another licence is stated on the immediately following page** this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# **Molecular Genetic Analysis of Hidradenitis Suppurativa (Acne Inversa)**

By

**Andrew Edward Pink**

This thesis is presented to the Faculty of Medicine, University of London  
for the degree of Doctor of Philosophy.

St. John's Institute of Dermatology

Division of Genetics and Molecular Medicine

Guy's King's and St Thomas' (GKT) School of Medicine

King's College London

July 2014

## **Declaration**

The work presented in this thesis is my own. All experiments, except where acknowledged in the text, were performed by me.

Andrew Edward Pink

## **Acknowledgements**

I owe a great debt of gratitude to Dr Michael Simpson who helped me to prepare my MRC grant application and supervised all of my laboratory work. He provided constant support and enthusiasm, championed successful submissions for further funding and taught me a great deal about drafting papers and indeed theses! I would like to thank my supervisors, Professors Richard Trembath and Jonathan Barker, who created the project, supported me through the funding application process, provided essential advice and direction to my work and acted as inspirational mentors. I am very grateful to Professor Catherine Smith, Robert Pleass, Naomi Hare, Karen Robertson, Kate Thornberry, Michelle Weston, Neil Green and Michael Duckworth for helping with ethics applications, patient recruitment and data collection and storage.

Dr Nemesha Desai and Professor Peter Mortimer were instrumental in devising and initiating this study. I would like to thank them for their encouragement, enthusiasm and most importantly for the invaluable clinical input that they provided throughout this project. They were integral to phenotyping patients for this research.

I could not have completed this work without the advice and assistance of multiple individuals in the laboratory. Mike Allen taught me countless techniques, provided valuable advice throughout and somehow found a solution to any problem! Dimitra Dafou taught me cell culture techniques and all the relevant skills required to perform the functional studies presented as part of this thesis. Wesley Woollard taught me the technique of exome sequencing. Alison Hills taught me how to perform multiplex ligation-dependent probe amplification. Isabella Tosi embedded skin specimens in paraffin blocks. Carl Hobbs very skilfully performed all of the immunohistochemistry experiments described in this thesis. Dr Oliver Holmes, Harvard Medical School, was



an invaluable collaborator in this research, providing the C100-Flag plasmid and the protocol required to perform the gamma-secretase assays. Due to problems optimising those assays he ultimately, and very kindly, performed the assays at the Institute for Neurological Disease, Harvard Medical School, Boston, USA. I would like to thank the Trembath laboratory group (particularly Francesca Capon and Laura Southgate) and the clinical fellows in the laboratory (Alex Navarini, Angela Tewari, Debra Josephs, Emma Benton, Joey Lai-Cheong, Alya Al-Wahib, Laura Proudfoot, Panos Karagiannis and Gaby Petrof) for their help and advice throughout this project. I would also like to thank Professor McGrath for his extremely valuable advice, encouragement and support over the last 4 years.

It has been a great honour and privilege to be able to take three and a half years out of clinical training to complete this project and that would not have been possible without funding from the National Institute for Health Research (NIHR), Medical Research Council (MRC), British Association of Dermatologists (BAD) and British Skin Foundation (BSF) (NIHR Academic Clinical Fellowship, MRC/ BAD/ BSF Clinical Research Training Fellowship), to whom I am extremely grateful. I would also like to thank all of the patients and their family members for the taking the time to take part in these studies.

It has been an exciting few years, both scientifically and at home. Our beautiful daughter, Lucy, arrived half way through this project. She is now an HS expert! I owe a great debt of gratitude to my wife, Katharine, who has been a source of constant support, enthusiasm and joy and has always shown a great interest in my work. She has made many sacrifices to allow me to complete this project for which I am extremely grateful. I would also like to thank my family, Mum, Dad and my brother Michael, for their never ending encouragement, advice, dedication and championing throughout my life, without which, none of this would have been possible.

## Abstract

Hidradenitis suppurativa (HS) is a chronic inflammatory dermatosis that presents with painful nodules, abscesses and sinus tracts in apocrine gland-bearing regions. Disease pathogenesis is poorly understood and there is a consequent paucity of effective treatment options. HS has been associated with smoking and obesity and can be inherited in an autosomal dominant (AD) manner. Heterozygous mutations in the gamma-secretase genes *NCSTN*, *PSEN1* and *PSENEN* have recently been reported in some Chinese multiplex kindreds with HS. The overall aim of this programme of research was to understand the genetic architecture of HS, the ultimate goal being to clarify the molecular mechanisms involved in disease pathogenesis. The principle objectives were four-fold: 1/ Explore the involvement of mutations in *NCSTN*, *PSEN1* and *PSENEN* in British familial cases of HS. 2/ Functionally characterise any identified mutations. 3/ Determine the contribution of mutations in *NCSTN*, *PSEN1* and *PSENEN* in the general disease population. 4/ Analyse the clinical phenotype of individuals harbouring gamma-secretase gene mutations.

Heterozygous gamma-secretase gene mutations were detected in affected individuals from two of seven British multiplex kindreds demonstrating AD inheritance. Mutations were detected in *NCSTN* (c.1125+1 G>A) and *PSENEN* (c.66\_67insG) which resulted in aberrant splicing (p.Glu333\_Gln367del) and a frameshift (p.Phe23ValfsX98) respectively. Both mutant transcripts appeared unstable and subject to decay in primary human fibroblasts. A corresponding reduction in respective protein expression in mutant versus wild type fibroblasts implied that haploinsufficiency of the gamma-secretase components may underlie the development of HS in those cases. This apparent haploinsufficiency did not however affect gamma-secretase endopeptidase or carboxypeptidase activity *in vitro*.

Forty eight individuals were sequentially recruited from a tertiary referral HS clinic and screened for variation in *NCSTN*, *PSENEN* and *PSEN1*. Three individuals harboured novel variants in *NCSTN* but no novel or rare (<1% population frequency) variation was detected in *PSEN1* or *PSENEN*. All three *NCSTN* variants were of uncertain pathogenicity. Mutations in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1* would therefore appear to underlie up to 7% of British cases of HS.

Phenotypic analysis of individuals with likely pathogenic gamma-secretase mutations revealed early onset, severe, widespread and treatment resistant disease often associated with atypical flexural pigmentation and prominent cystic changes in atypical areas.

Given that gamma-secretase gene mutations only appear to underlie a small proportion of HS cases it was hypothesised that there is further genetic heterogeneity underlying HS. Unexplained familial cases and unrelated but carefully phenotyped and subgrouped individuals were therefore assessed using a combination of traditional and next generation gene mapping techniques. No pathogenic mutations in additional genes were identified.

The identification of gamma-secretase gene mutations in HS informs disease pathogenesis, potentially implicates the gamma-secretase-Notch signalling axis in disease development, provides a platform for ongoing functional studies and will hopefully facilitate the identification of new therapeutic targets in this debilitating disease. Ongoing genetic and functional investigation is now required to confirm the presence of further genetic heterogeneity and to establish the relevance of gamma-secretase-Notch signalling in the wider patient cohort.

## Abbreviations

AA	Amino acid
A $\beta$	Amyloid beta
AD	Autosomal dominant
AICD	APP intracellular domain
APP	Amyloid precursor protein
ASSED	Automated Splice Site and Exon Definition Analysis
BAD	British Association of Dermatologists
BCA	Bicinchoninic acid assay
BMI	Body Mass Index
BMZ	Basement membrane zone
BSF	British Skin Foundation
cDNA	Complementary DNA
CNV	Copy number variation
CRF	Clinical Research Form
CTF	C terminal fragment
DAMPS	Damage associated molecular patterns
DDD	Dowling-Degos disease
DLQI	Dermatology Life Quality Index
DNA	Deoxyribonucleic acid
ECL	Enhanced chemiluminescence
EGFR	Endocrine growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
EVS	Exome variant server
FAD	Familial Alzheimer's disease
FBS	Fetal bovine serum
FISH	Flourescence in situ hybridisation
GCP	Good Clinical Practice course
HDL	High-density lipoprotein
HLA	Human leukocyte antigen
HS	Hidradenitis suppurativa
HSS	Hidradenitis Severity Score
IHC	Immunohistochemistry
KID	Keratitis, ichthyosis and deafness (syndrome)
LB	Luria Bertani medium
LOD	Logarithm of odds
MEF	Mouse embryonic fibroblasts
MLPA	Multiplex ligation-dependent probe amplification
MRC	Medical Research Council
Mrna	Messenger RNA
N $\beta$	Notch $\beta$ -peptide
NEXT	Notch extra-cellular truncation
NHS	National Health Service
NICD	Notch intracellular domain
NIH	National Institutes of Health

NIHR	National Institute for Health Research
NK	Natural killer (cells)
NMD	Nonsense mediated decay
NSAID	Non-steroidal anti-inflammatory drug
NTF	N terminal fragment
PAMPS	Pathogen associated molecular patterns
PAPA	Pyogenic arthritis, pyoderma gangrenosum and acne
PAPASH	Pyogenic arthritis, pyoderma gangrenosum, acne, hidradenitis suppurativa
PASH	Pyoderma gangrenosum, acne and suppurative hidradenitis
PBS	Phosphate buffered saline
PC	L- $\alpha$ -phosphatidylethanolamine chloroform
PCR	Polymerase chain reaction
PE	L- $\alpha$ -phosphatidylethanolamine chloroform
PSH	Presenilin homologue
PUVA	Psoralen ultraviolet A
RIP	Regulated intramembranous proteolysis
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SCC	Squamous cell carcinoma
SD	Standard deviation
SE	Standard error
SiRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
TLR	Toll-like receptor
TM	Trans-membrane
TNF	Tumour necrosis factor
US	United States
USA	United States of America
UV	Ultraviolet
UVA	Ultraviolet A
WES	Whole exome sequencing
WT	Wild type

# Table of contents

<b>DECLARATION.....</b>	<b>2</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>3</b>
<b>ABSTRACT.....</b>	<b>5</b>
<b>ABBREVIATIONS.....</b>	<b>7</b>
<b>TABLE OF FIGURES.....</b>	<b>21</b>
<b>TABLE OF TABLES.....</b>	<b>27</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>30</b>
1.1 HIDRADENITIS SUPPURATIVA.....	30
1.1.1 Disease definition .....	30
1.1.2 Epidemiology.....	31
1.1.3 Clinical presentation .....	32
1.1.4 Histopathology .....	33
1.1.5 Phenotypic classification .....	34
1.1.6 Clinical diagnosis and staging.....	35
1.1.7 The impact of the disease on quality of life.....	36
1.1.8 Treatment of hidradenitis suppurativa .....	37
1.1.9 Conditions associated with hidradenitis suppurativa .....	38
1.2 PATHOGENESIS OF HIDRADENITIS SUPPURATIVA.....	40
1.2.1 Aberrant immune responses in HS .....	41
1.2.2 Hormonal influence .....	45
1.2.3 Bacterial infection .....	46
1.2.4 Obesity and smoking.....	47
1.2.5 The genetics of hidradenitis suppurativa.....	48

1.2.6	<i>Mutations in the gamma-secretase genes .....</i>	48
1.3	<b>GAMMA-SECRETASE .....</b>	50
1.3.1	<i>The composition of the gamma-secretase complex .....</i>	52
1.3.2	<i>The assembly of the gamma-secretase complex.....</i>	63
1.3.3	<i>The structure of the gamma-secretase complex .....</i>	64
1.3.4	<i>Proteins that associate with the gamma-secretase complex.....</i>	66
1.3.5	<i>Gamma-secretase processing of APP.....</i>	70
1.3.6	<i>Gamma-secretase processing of Notch.....</i>	72
1.3.7	<i>PSEN1 and PSEN2 mutations reported in familial Alzheimer's disease(FAD).....</i>	74
1.4	<b>AIMS AND OBJECTIVES .....</b>	77
 <b>CHAPTER 2: MATERIALS AND METHODS .....</b>		<b>79</b>
2.1	<b>PATIENT AND CONTROL COHORTS.....</b>	79
2.1.1	<i>Regulatory and ethical approval.....</i>	79
2.1.2	<i>Clinical definition of Hidradenitis Suppurativa.....</i>	79
2.1.3	<i>Patient recruitment .....</i>	80
2.1.4	<i>Recruitment of family members.....</i>	80
2.1.5	<i>Recruitment of healthy volunteer subjects (controls).....</i>	81
2.1.6	<i>Recruitment of patients with nodulocystic acne .....</i>	81
2.2	<b>ANTIBODIES USED .....</b>	83
2.3	<b>BUFFERS USED.....</b>	85
2.4	<b>MOLECULAR BIOLOGY TECHNIQUES .....</b>	87
2.4.1	<i>DNA extraction from blood (leucocytes) .....</i>	87
2.4.2	<i>DNA extraction from saliva .....</i>	87
2.4.3	<i>Polymerase chain reaction (PCR).....</i>	88

2.4.4	<i>Analysis of PCR products by agarose gel electrophoresis .....</i>	89
2.4.5	<i>Direct nucleotide sequencing .....</i>	89
2.4.6	<i>Genotyping using Exome chip.....</i>	90
2.4.7	<i>Exome chip data analysis to define regions of linkage within pedigrees ...</i>	91
2.4.8	<i>Whole exome sequencing method.....</i>	91
2.4.9	<i>Whole exome sequencing data preparation .....</i>	92
2.4.10	<i>RNA isolation from whole blood .....</i>	92
2.4.11	<i>RNA isolation from cells .....</i>	93
2.4.12	<i>RNA interference .....</i>	95
2.4.13	<i>Verification of the specificity of an NCSTN antibody by RNA interference .....</i>	96
2.4.14	<i>Reverse transcription polymerase chain reaction (RT-PCR) .....</i>	96
2.4.15	<i>cDNA PCR and direct sequencing .....</i>	97
2.4.16	<i>Gene expression assays .....</i>	98
2.4.17	<i>Multiplex Ligation-dependent Probe Amplification (MLPA) .....</i>	100
2.4.18	<i>Preparation of Luria Bertani (LB) medium and 2 x LB media .....</i>	102
2.4.19	<i>Preparation of C100-Flag pET-21b (-) plasmid.....</i>	103
2.4.20	<i>Transformation of NEB 5-alpha competent E Coli with C100-Flag pET-21b (-) plasmid.....</i>	103
2.4.21	<i>Plasmid purification .....</i>	104
2.4.22	<i>Transformation of E-Coli BL21 DE3 cells with C100Flag – pET-21b plasmid.....</i>	104
2.5	<b>IMMUNOHISTOCHEMISTRY.....</b>	106
2.5.1	<i>Embedding of skin biopsies in paraffin.....</i>	106
2.5.2	<i>Sectioning and staining paraffin embedded skin.....</i>	106
2.6	<b>CELL CULTURE.....</b>	108



2.6.1	<i>Isolation of primary fibroblasts from skin biopsies.....</i>	108
2.6.2	<i>Passage of primary human fibroblasts .....</i>	108
2.6.3	<i>Maintenance of primary human fibroblasts.....</i>	109
2.6.4	<i>Maintenance of S20 cell line.....</i>	109
2.6.5	<i>Storing cells in liquid nitrogen.....</i>	109
2.7	PROTEIN ANALYSIS .....	111
2.7.1	<i>Isolation of total protein from primary human fibroblasts in culture .....</i>	111
2.7.2	<i>Isolation and solubilisation of cell membrane proteins (including gamma-secretase microsomes) from cells in culture.....</i>	111
2.7.3	<i>Protein quantification.....</i>	112
2.7.4	<i>Immunoblotting Nicastrin and PEN-2.....</i>	114
2.7.5	<i>Immunoblotting <math>\beta</math>-actin.....</i>	115
2.7.6	<i>Immunoblotting Flag-tagged proteins.....</i>	116
2.7.7	<i>Analysis of western blots.....</i>	116
2.8	CELL BASED ASSAYS .....	117
2.8.1	<i>Addition of cycloheximide to assess for nonsense mediated decay.....</i>	117
2.9	GAMMA-SECRETASE ENZYME ACTIVITY ASSAYS .....	118
2.9.1	<i>Purification of C100-Flag substrate.....</i>	118
2.9.2	<i>Gamma-secretase activity assay .....</i>	118
2.10	STATISTICAL ANALYSIS.....	121

## **CHAPTER 3: RESULTS: MUTATIONS IN THE GAMMA-SECRETASE**

### **GENES *PSENEN* AND *NCSTN* UNDERLIE SOME**

### **FAMILIAL FORMS OF HIDRADENITIS SUPPURATIVA ..... 122**

3.1	INTRODUCTION .....	122
3.1.1	<i>Hypothesis and objectives.....</i>	122

3.2	PATIENT COHORT AND SAMPLES .....	123
3.3	IDENTIFICATION OF A NOVEL HETEROZYGOUS SINGLE BASE INSERTION IN <i>PSENEN</i> (c.66_67INSG) .....	128
3.4	IDENTIFICATION OF A NOVEL HETEROZYGOUS SINGLE NUCLEOTIDE SUBSTITUTION IN THE <i>NCSTN</i> EXON 9/INTRON9 DONOR SPLICE SITE (c.1125+1 G>A).....	130
3.5	NO LARGE SCALE DELETIONS OR DUPLICATIONS IN <i>NCSTN</i> , <i>PSENEN</i> AND <i>PSEN1</i> WERE DETECTED IN ANY OF THE SEVEN KINDREDS .....	131
3.5.1	<i>Methodological approaches to detecting copy number variation (CNV).....</i>	131
3.5.2	<i>Multiplex ligation-dependent probe amplification (MLPA) .....</i>	132
3.5.3	<i>No whole exon or gene deletions or duplications were detected in the gamma-secretase genes NCSTN, PSEN1 and PSENEN in this HS cohort.....</i>	132
3.6	DISCUSSION.....	134

<b>CHAPTER 4: RESULTS: FUNCTIONAL CHARACTERISATION OF THE GAMMA-SECRETASE MUTATIONS IDENTIFIED IN <i>PSENEN</i> AND <i>NCSTN</i> IN FAMILIAL HS .....</b>		<b>135</b>
4.1	INTRODUCTION .....	135
4.1.1	<i>Aims and objectives .....</i>	135
4.2	PATIENT COHORT AND SAMPLES .....	136
4.3	<i>NCSTN</i> TRANSCRIPT AND <i>NCSTN</i> PROTEIN ANALYSIS IN AN INDIVIDUAL HARBOURING THE HETEROZYGOUS <i>NCSTN</i> SPLICE SITE SUBSTITUTION <i>NCSTN</i> c.1125+1 G>A.....	137
4.3.1	<i>NCSTN c. 1125+1 G&gt;A results in the skipping of NCSTN exon 9 .....</i>	137

4.3.2	<i>NCSTN transcript abundance is reduced as a consequence of NCSTN c.1125+1 G&gt;A.....</i>	141
4.3.3	<i>The mutant NCSTN transcript is subject to decay .....</i>	142
4.3.4	<i>Nicastrin protein expression is reduced in mutant fibroblasts versus control fibroblasts.....</i>	146
4.4	<i>PSENEN TRANSCRIPT AND PSENEN PROTEIN ANALYSIS IN AN INDIVIDUAL HARBOURING THE HETEROZYGOUS PSENEN INSERTION PSENEN C.66_67INSG .....</i>	152
4.4.1	<i>PSENEN c.66_67insG results in a frameshift and is predicted to result in an altered and lengthened protein product .....</i>	152
4.4.2	<i>PSENEN transcript abundance is reduced as a consequence of PSENEN c.66_67insG .....</i>	154
4.4.3	<i>The mechanism underlying the apparent reduction in PSENEN transcript abundance is unclear.....</i>	155
4.4.4	<i>PEN-2 protein expression was reduced in mutant fibroblasts compared to controls .....</i>	158
4.5	<i>THE EFFECT OF NCSTN C.1125+1 G&gt;A AND PSENEN C.66_67INSG MUTATIONS ON GAMMA-SECRETASE ENZYME ACTIVITY IN VITRO.....</i>	161
4.5.1	<i>Methodological approaches to evaluating gamma-secretase activity in vitro.....</i>	161
4.5.2	<i>Analysis of in vitro gamma-secretase activity using an amyloid precursor protein-based assay (C100F substrate).....</i>	161
4.5.3	<i>Production of the C100-Flag enzyme substrate .....</i>	164
4.5.4	<i>Initiation and maintenance of an S-20 cell line .....</i>	167
4.5.5	<i>Purification and solubilisation of membrane proteins from primary human fibroblasts and S20 cells.....</i>	167

4.5.6	<i>Optimisation of the gamma-secretase activity assay.....</i>	171
4.5.7	<i>Gamma-secretase activity assays.....</i>	177
4.5.8	<i>Analysis of gamma-secretase enzyme activity in mutant and control primary human fibroblasts.....</i>	179
4.5.9	<i>In contrast to total cell fractions, NCSTN and PEN-2 expression in solubilised cell membrane fractions is equal in mutant and control fibroblasts.....</i>	185
4.6	IMMUNOHISTOCHEMICAL ANALYSIS OF NCSTN AND PEN-2 EXPRESSION IN THE SKIN .....	189
4.6.1	<i>Primary antibody optimisation.....</i>	189
4.6.2	<i>Gamma-secretase is concentrated in the epidermis, hair follicle, apocrine gland and sebaceous gland in axillary skin harvested from healthy volunteers.....</i>	190
4.6.3	<i>NCSTN and PEN-2 are expressed in the epidermis, hair follicle, apocrine gland, sebaceous gland and inflammatory infiltrate in axillary skin harvested from patients with NCSTN c.1125+1 G&gt;A and PSENEN c.66_67insG mutations .....</i>	192
4.6.4	<i>There was no discernible difference in epidermal Nicastrin and PEN-2 staining intensity in axillary skin harvested from patients with mutations versus healthy volunteers.....</i>	197
4.7	DISCUSSION.....	201
4.7.1	<i>Both mutations appear to result in haploinsufficiency but the respective protein expression at the cell membrane remains unaffected.....</i>	202
4.7.2	<i>These mutations are not associated with a significant difference in gamma-secretase complex number, structure or activity .....</i>	203

4.7.3	<i>Cutaneous expression of the gamma-secretase components correlates with the histopathological changes observed in HS, supporting a role for gamma-secretase mutations in disease pathogenesis</i> .....	205
4.7.4	<i>Conclusions</i> .....	206

## CHAPTER 5: RESULTS: MUTATIONS IN THE GAMMA-SECRETASE

### GENES *NCSTN*, *PSENEN* AND *PSEN1* UNDERLIE A

### MINORITY OF CASES OF HIDRADENITIS

### SUPPURATIVA..... 210

5.1	INTRODUCTION .....	210
5.1.1	<i>Aims and objectives</i> .....	210
5.2	PATIENT COHORT .....	211
5.3	THREE OF FORTY EIGHT INDIVIDUALS HARBOURED NOVEL VARIANTS IN THE GAMMA-SECRETASE GENES <i>NCSTN</i> , <i>PSEN1</i> AND <i>PSENEN</i> .....	213
5.4	IDENTIFICATION OF A NOVEL MISSENSE VARIANT IN EXON 5 OF <i>NCSTN</i> (C.553 G>A, P.ASP185ASN) .....	215
5.5	IDENTIFICATION OF A NOVEL HETEROZYGOUS SUBSTITUTION IN THE EXON 8/ INTRON 8 SPLICE SITE OF <i>NCSTN</i> (C.996+7G>A).....	218
5.6	IDENTIFICATION OF A NOVEL HETEROZYGOUS SUBSTITUTION IN THE <i>NCSTN</i> EXON 9/ INTRON 9 SPLICE SITE OF <i>NCSTN</i> (C.1101+10 A>G) .....	220
5.7	NO LARGE SCALE DELETIONS OR DUPLICATIONS DETECTED IN <i>NCSTN</i> , <i>PSENEN</i> AND <i>PSEN1</i> .....	222
5.8	A PILOT STUDY TO INVESTIGATE THE ROLE OF <i>NCSTN</i> MUTATIONS IN NODULOCYSTIC ACNE VULGARIS REVEALED A NOVEL MISSENSE VARIANT ( <i>NCSTN</i> C.1315 G>A, P.VAL439ILE) .....	223
5.8.1	<i>Patient cohort</i> .....	223

5.8.2	<i>A novel missense variant (NCSTN C.1315 G&gt;A, p.Val439Ile) was detected in one individual with nodulocystic acne vulgaris</i>	223
5.9	DISCUSSION	225

## CHAPTER 6: INVESTIGATION TO DETERMINE FURTHER DISEASE

	<b>CAUSING GENES IN HS</b>	<b>228</b>
6.1	INTRODUCTION	228
6.1.1	<i>Hypothesis and objectives</i>	228
6.1.2	<i>Methodological considerations</i>	228
6.2	GENETIC METHODS AND PATIENT SELECTION	235
6.2.1	<i>WES, combined with linkage analysis in familial cases, was deemed an appropriate method by which to further interrogate the genetic architecture of HS</i>	235
6.2.2	<i>Whole exome sequencing</i>	235
6.2.3	<i>Genotyping of individuals from the multiplex kindreds and linkage analysis</i>	240
6.2.4	<i>Data analysis strategy</i>	241
6.3	RESULTS	246
6.3.1	<i>Exome sequencing coverage</i>	246
6.3.2	<i>Analysis of the gamma-secretase genes in all individuals subjected to exome sequencing</i>	247
6.3.3	<i>Familial analysis</i>	251
6.3.4	<i>Analysis of phenotypically grouped, unrelated individuals with HS</i>	254
6.4	DISCUSSION	258
6.4.1	<i>Three novel mutations in NCSTN were identified through WES</i>	258
6.4.2	<i>Familial studies identify potential candidate variants</i>	258

6.4.3	<i>Exclusion of genetic homogeneity outside of the gamma-secretase genes .....</i>	259
6.4.4	<i>Analysis strategies employed to interpret WES data.....</i>	259
6.4.5	<i>Potential analysis strategies that could be employed in the future interrogation of these data.....</i>	261
6.4.6	<i>Conclusion .....</i>	264

## **CHAPTER 7: PHENOTYPIC ANALYSIS OF INDIVIDUALS**

### **HARBOURING MUTATIONS IN THE GAMMA-SECRETASE GENES *NCSTN* AND *PSENEN* ..... 265**

7.1	INTRODUCTION AND AIMS.....	265
7.1.1	<i>Aims and objectives .....</i>	265
7.2	PATIENT COHORT .....	266
7.3	CLINICAL PHENOTYPE OF PATIENTS HARBOURING GAMMA-SECRETASE GENE MUTATIONS .....	268
7.4	DISCUSSION .....	271

## **CHAPTER 8: CONCLUSIONS AND FUTURE WORK..... 276**

8.1	INTRODUCTION .....	276
8.2	GAMMA-SECRETASE GENE MUTATIONS IN HS .....	277
8.2.1	<i>A loss of function of the gamma-secretase genes appears to underlie HS .....</i>	279
8.2.2	<i>Cell membrane expression of the gamma-secretase components appears tightly regulated and gamma-secretase enzyme activity was not significantly impaired in mutant fibroblasts in vitro .....</i>	280

8.2.3	<i>Immunohistochemical studies of human skin and in vivo gamma-secretase gene mutagenesis animal models support a role for gamma-secretase mutations in HS</i> .....	281
8.2.4	<i>Gamma-secretase mutations may result in HS by affecting Notch signalling</i> .....	283
8.2.5	<i>Gamma-secretase mutations provide an insight in to the pathogenic mechanisms underlying HS</i> .....	284
8.2.6	<i>Gamma-secretase mutations underlie a minority of HS cases in the UK</i> .....	285
8.2.7	<i>A gamma-secretase mutation-related clinical phenotype is emerging which represents the first genetically defined phenotypic sub-group of HS</i> .....	286
8.2.8	<i>There is no evidence to support an association between HS and FAD....</i>	287
8.3	THE FUTURE OF GAMMA-SECRETASE RELATED RESEARCH IN HS.....	289
8.3.1	<i>Future functional characterisation of the gamma-secretase mutations identified in HS</i> .....	289
8.3.2	<i>The translational relevance of the above functional studies is partially dependent upon the prevalence of gamma-secretase gene mutations in the general disease population</i> .....	293
8.3.3	<i>Functional characterisation of the gamma-secretase gene mutations identified in HS may provide a novel insight into general gamma-secretase enzyme complex function</i> .....	294
8.3.4	<i>Further investigations are required to determine if there is any association between HS and FAD</i> .....	295
8.3.5	<i>Refinement of the gamma-secretase related clinical phenotype</i> .....	295
8.4	ONGOING GENETIC INVESTIGATION OF HS .....	297



8.4.1	<i>Further genetic heterogeneity in HS .....</i>	297
8.4.2	<i>A combination of traditional and next generation methods of genetic interrogation identified further potential candidate genes in HS and excluded genetic homogeneity outside of the gamma-secretase genes ....</i>	297
8.4.3	<i>The future of genetic studies in HS.....</i>	298
8.5	CONCLUDING REMARKS .....	301
<b>REFERENCES.....</b>		<b>302</b>
<b>APPENDIX 1.....</b>		<b>330</b>
<b>APPENDIX 2.....</b>		<b>332</b>
<b>APPENDIX 3.....</b>		<b>377</b>
<b>APPENDIX 4.....</b>		<b>378</b>
<b>APPENDIX 5.....</b>		<b>380</b>

## Table of figures

Figure 1.1. The clinical appearance of HS. ....	33
Figure 1.2. The current histopathological model of HS.....	41
Figure 1.3. The key roles of gamma-secretase in cellular functioning. ....	52
Figure 1.4. The gamma-secretase complex. ....	53
Figure 1.5. The 3D structure of the archael presenilin homologue (PSH).....	65
Figure 1.6. The predicted structure of the gamma-secretase complex. ....	66
Figure 1.7. $\beta$ -secretase/ $\gamma$ -secretase processing of amyloid precursor protein (APP). ...	71
Figure 1.8. $\alpha$ -secretase/ $\gamma$ -secretase processing of amyloid precursor protein (APP). ...	72
Figure 1.9. Gamma-secretase processing of Notch receptors. ....	74
Figure 2.1. Example of MLPA results. ....	102
Figure 2.2. Example of a standard curve plotted for albumin protein standards.....	114
Figure 2.3. An example western blot generated from a gamma-secretase activity assay. ....	120
Figure 3.1. Family pedigrees recruited for genetic analysis. ....	124
Figure 3.2. Family pedigrees recruited for genetic analysis. ....	125
Figure 3.3. Family pedigrees recruited for genetic analysis. ....	126
Figure 3.4. The <i>PSENEN</i> c.66_67insG mutation was detected in affected individuals from Pedigree 3.....	129
Figure 3.5. The <i>NCSTN</i> c.1125+1 G>A mutation was detected in an affected individual from pedigree 4. ....	130
Figure 3.6. A plot of the normalised MLPA peak ratios in one individual with HS (probes designed to assess copy number of the exons of <i>PSEN1</i> and <i>PSENEN</i> ).....	133
Figure 4.1. Primary human fibroblasts grown in culture. ....	137

Figure 4.2. Agarose electrophoresis and Sanger sequencing of the full length <i>NCSTN</i> transcript in the patient harbouring the heterozygous <i>NCSTN</i> 1125+1 G>A mutation.....	139
Figure 4.3. The consequences of the <i>NCSTN</i> c.1125+1 G>A mutation on the <i>NCSTN</i> transcript and protein. ....	140
Figure 4.4. Relative <i>NCSTN</i> transcript abundance in the individual harbouring <i>NCSTN</i> c.1125+1 G>A versus healthy volunteers (controls).....	142
Figure 4.5. <i>NCSTN</i> transcript abundance within mutant and control fibroblasts following incubation with cycloheximide. ....	144
Figure 4.6. Predicted RNA secondary structure of the WT and mutant (c.1125+1 G>A) <i>NCSTN</i> transcript. ....	145
Figure 4.7. <i>NCSTN</i> transcript abundance and NCSTN protein expression in primary human fibroblasts following SiRNA knockdown.....	149
Figure 4.8. Immunoblot of NCSTN in mutant and control primary human fibroblasts. ....	151
Figure 4.9. <i>PSENEN</i> cDNA sequence in an individual harbouring the heterozygous c.66_67insG. ....	152
Figure 4.10. The predicted <i>PSENEN</i> c.66_67insG mutant transcript and protein. ....	153
Figure 4.11. Relative <i>PSENEN</i> transcript abundance in mutant and control primary human fibroblasts.....	155
Figure 4.12. Predicted RNA secondary structure of the WT and mutant (c.66_67insG) <i>PSENEN</i> transcript.....	156
Figure 4.13. <i>PSENEN</i> transcript abundance in mutant fibroblasts following incubation with cycloheximide. ....	158
Figure 4.14. Immunoblot of PEN-2 in protein samples harvested from mutant and control primary human fibroblasts. ....	160

Figure 4.15. $\beta$ -secretase and $\gamma$ -secretase cleavage and processing of amyloid precursor protein in the cell membrane. ....	162
Figure 4.16. Principle of the C100-Flag APP based gamma-secretase activity assay..	163
Figure 4.17. Protein electrophoresis of pre- and post-IPTG induced cell lysate samples from BL21 (DE3) E.coli cells expressing the C100-Flag pET-21b plasmid.....	165
Figure 4.18. Electrophoresis and InstantBlue staining of samples generated during the purification of C100-Flag. ....	166
Figure 4.19. Immunoblot of purified C100-Flag samples using a primary anti-Flag antibody. ....	166
Figure 4.20. S20 cells in culture. ....	167
Figure 4.21. Immunoblotting of NCSTN and PEN-2 in solubilised membrane preparations from S20 cells and primary human fibroblasts. ....	169
Figure 4.22. Standard curve plotted using protein concentration and 562nm light absorbance readings for protein standards. ....	170
Figure 4.23. Immunoblot showing the results of a gamma-secretase activity assay. ...	171
Figure 4.24. Immunoblot of a gamma-secretase activity assay reaction performed using enzyme derived from S20 cells. ....	173
Figure 4.25. Immunoblot of enzyme activity assays using variable concentrations of S20 solubilised membrane protein.....	174
Figure 4.26. Immunoblot of gamma-secretase assays using variable amounts of C100-Flag substrate. ....	175
Figure 4.27. Immunoblot of gamma-secretase enzyme activity assays using different batches of and no cholesterol. ....	176
Figure 4.28 A representative gamma-secretase activity assay immunoblot. ....	178

Figure 4.29. PSEN1-CTF expression in solubilised membrane preparations harvested from mutant and control fibroblasts. ....	180
Figure 4.30. Crude gamma-secretase activity per mg of total solubilised membrane protein expressed as a percentage of the positive control (S20 cells). ....	181
Figure 4.31. Gamma-secretase enzyme activity per complex expressed as a percentage of the positive control (S20 cell line). ....	182
Figure 4.32. PSEN1 CTF/ PEN-2 ratio for each sample expressed as a percentage of the ratio observed in the positive control S20 cells. ....	184
Figure 4.33. A $\beta$ 42: A $\beta$ 40 ratios expressed as a percentage of the ratio derived from S20 cells. ....	185
Figure 4.34. NCSTN expression in solubilised cell membrane fractions harvested from mutant and control fibroblasts. ....	187
Figure 4.35. PEN-2 expression in solubilised cell membrane fractions harvested from mutant and control fibroblasts. ....	188
Figure 4.36. Nicastrin and PEN-2 staining in human breast cancer and healthy skin. .	190
Figure 4.37. NCSTN and PEN-2 expression in healthy volunteer axillary skin. ....	191
Figure 4.38 Close up of NCSTN and PEN-2 staining in the epidermis. ....	192
Figure 4.39. Cutaneous NCSTN expression in affected axillary skin from the individual harbouring the <i>NCSTN</i> c.1125+1 G>A mutation. ....	193
Figure 4.40. Cutaneous PEN-2 staining in affected axillary skin from the individual harbouring the <i>PSENEN</i> c.66_67insG mutation. ....	194
Figure 4.41. Immunohistochemical staining of the inflammatory infiltrate observed in axillary skin harvested from the patient harbouring the <i>NCSTN</i> c.1125+1 G>A mutation. ....	195

Figure 4.42. Immunohistochemical staining of the inflammatory infiltrate observed in axillary skin harvested from the patient with the <i>PSENEN</i> 66_67insG mutation. ....	196
Figure 4.43. A comparison of NCSTN epidermal staining in axillary skin harvested from the patient with an <i>NCSTN</i> c.1125+1 G>A mutation and control subjects. ....	198
Figure 4.44. A comparison of PEN-2 epidermal staining in axillary skin harvested from the patient harbouring the <i>PSENEN</i> c.66_67insG mutation and healthy volunteers. ....	199
Figure 4.45. A comparison of PEN-2 epidermal staining in para-axillary skin harvested from the individual harbouring <i>PSENEN</i> c.66_67insG and healthy volunteers. ....	200
Figure 5.1. Three individuals with novel variants in <i>NCSTN</i> . ....	214
Figure 5.2. Conservation of the aspartic acid residue (D185) affected by the <i>NCSTN</i> c.553 G>A variant. ....	215
Figure 5.3. Sequence and agarose electrophoresis of the full length <i>NCSTN</i> transcript in the individual with an <i>NCSTN</i> c.553 G>A variant. ....	217
Figure 5.4. Sequence and agarose electrophoresis of the full length <i>NCSTN</i> transcript in the individual harbouring an <i>NCSTN</i> c.996+7 G>A variant. ....	219
Figure 5.5. Sequence and agarose electrophoresis of the full length <i>NCSTN</i> transcript in the individual with an <i>NCSTN</i> c.1101+10 A>G variant. ....	221
Figure 5.6. A plot of the normalised MLPA peak ratios derived in one individual with HS (probes designed to assess copy number of the exons of <i>NCSTN</i> ). ....	222
Figure 5.7. A novel heterozygous substitution in <i>NCSTN</i> in a patient with nodulocystic acne vulgaris. ....	224

Figure 5.8. Conservation of the NCSTN Val439 residue. ....	224
Figure 6.1. Multiplex kindreds studied using WES.....	237
Figure 6.2. The algorithm used to select individuals within the sinus tract and nodular disease cohorts for WES.....	239
Figure 6.3. A figure showing the individuals within each pedigree that were genotyped using the Illumina Infinium HumanExome 12v1.1 chip to generate linkage data. ....	241
Figure 6.4. Algorithm for filtering coding variants or variants detected within 10bp of splice site boundaries identified through WES.....	243
Figure 6.5. Algorithm for filtering CNV identified by WES. ....	244
Figure 6.6. Algorithm detailing how data from individuals within the familial and unrelated cohorts was generated, collated and filtered prior to ongoing analysis. ....	245
Figure 6.7. Variant statistics for all individuals subjected to WES. ....	247
Figure 6.8. Sanger sequencing validated the WES results and confirmed the presence of novel heterozygous variants in <i>NCSTN</i> in three individuals. .	249
Figure 6.9. A proposed detailed variant analysis algorithm.....	262
Figure 7.1. Clinical images of individuals with likely pathogenic gamma-secretase mutations.....	270
Figure 7.2 A diagram representing the different sub-types of HS that have now been described. ....	275
Figure 8.1. The location of all gamma-secretase gene mutations now reported in HS. ....	280

## Table of tables

Table 1.1. Proteins that have been shown to interact with PSEN1 in two independent studies .....	69
Table 2.1. Antibodies used .....	83
Table 2.2. In-house buffers.....	85
Table 2.3. PCR reaction mix .....	88
Table 2.4. cDNA reverse transcription mix. ....	96
Table 2.5. cDNA primers used for sequencing <i>NCSTN</i> and <i>PSENEN</i> cDNA .....	97
Table 2.6. PCR reaction mix used to amplify the full length <i>NCSTN</i> transcript.....	98
Table 2.7. Primers used to sequence the full length <i>NCSTN</i> transcript .....	98
Table 2.8. cDNA PCR reaction mix .....	99
Table 2.9. Taqman probes used to study <i>NCSTN</i> and <i>PSENEN</i> transcript expression.....	99
Table 2.10. Table demonstrating variation in light absorbance with protein standards.....	113
Table 4.1. Demographics and clinical details of the two mutation-positive patients and three healthy volunteer subjects. ....	136
Table 4.2. Number of cells harvested and amount of 1% CHAPSO used to extract solubilised membrane proteins.....	168
Table 4.3. Concentration of the solubilised cell membrane preparations as calculated using BCA assays. ....	170
Table 5.1. Population and clinical data on the 48 HS patients recruited to the study .....	212



Table 6.1. A table showing the phenotypic subgroups analysed via WES and the number of individuals within each group. ....	240
Table 6.2. Coverage statistics for all individuals subjected to WES .....	246
Table 6.3. All variants with a population frequency < 5% detected in the gamma-secretase genes in the cohort of individuals subjected to WES. ....	248
Table 6.4. A table showing the number of potentially disease-causing variants identified by analyses 1 and 2 in pedigrees 1-6. ....	253
Table 6.5. A table showing the distribution in the number of genes in which novel genetic variation was common to X number of individuals within the HS/ arthritis cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene). ....	255
Table 6.6. A table showing the distribution in the number of genes in which rare (<1% population frequency) genetic variation was common to X number of individuals within the HS/ arthritis cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene). ....	255
Table 6.7. A table showing the distribution in the number of genes in which novel genetic variation was common to X number of individuals within the sinus tract disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene). ....	256
Table 6.8. A table showing the distribution in the number of genes in which rare genetic variation (<1% population frequency) was common to X number of individuals within the HS/ arthritis cohort (variation	

did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).....	256
Table 6.9. A table showing the distribution in the number of genes in which novel variation was common to X number of individuals within the nodular disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene). ....	257
Table 6.10. A table showing the distribution in the number of genes in which rare genetic variation (<1% population frequency) was common to X number of individuals within the nodular disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).....	257
Table 7.1. All of the likely pathogenic gamma-secretase gene variants identified in this study. ....	267
Table 7.2. A table summarising the clinically distinct characteristics observed in individuals harbouring gamma-secretase gene mutations.....	274
Table 8.1. All gamma-secretase mutations now reported in HS. ....	278

## **Chapter 1: Introduction**

### **1.1 Hidradenitis suppurativa**

Hidradenitis suppurativa (HS, acne inversa, OMIM 142690) was first described by Alfred Velpeau, a distinguished French surgeon, in 1839<sup>1</sup>. He described the formation of superficial abscesses in the axillae, sub-mammary folds and the perianal region. A surgical colleague, Aristide Verneuil, documented further cases in 1854 and, believing that it was a disease of sweat glands, proposed the name “hidrosadenite phlegmoneuse”<sup>2</sup>. Despite this, and as was tradition at the time, the condition was frequently referred to as Velpeau’s or Verneuil’s disease in the literature. As knowledge of sweat gland physiology evolved it was hypothesised that HS was primarily a disease of the apocrine gland<sup>3</sup>. This was supported by the presence of a peri-glandular inflammatory infiltrate in affected skin and the anatomical distribution of the disease (mainly but not exclusively affecting apocrine-gland bearing areas)<sup>4; 5</sup>. Further work demonstrating that physical occlusion of axillary skin induces apocrine duct plugging and nodule formation further supported this theory<sup>5</sup>. More recent histo-pathological evidence would however implicate the hair follicle rather than the apocrine gland in HS pathogenesis, suggesting that follicular plugging is the primary histopathological mechanism underlying HS<sup>6; 7</sup>. This has led some to question the accuracy of the term hidradenitis (inflammation of sweat glands) in describing this disease and an alternative name, “acne inversa”, has recently been proposed but is rarely used<sup>8</sup>.

#### **1.1.1 Disease definition**

A consensus definition for hidradenitis suppurativa was devised at the Second International HS Research Symposium, San Francisco, 2009:

“HS is a chronic, inflammatory, recurrent, debilitating, skin follicular disease that usually presents after puberty with painful deep seated, inflamed lesions in the apocrine

gland-bearing areas of the body, most commonly, the axillary, inguinal and anogenital regions.”

### **1.1.2 Epidemiology**

HS develops at an average age of 22.1 years (usually post-pubertal but can occur at any age) and persists for an average of 18.8 years<sup>9; 10</sup>. There is a female predominance, with reported female to male ratios as high as 5:1<sup>10</sup>. The 1 year prevalence is estimated to be 1-4%<sup>11; 12</sup> in European populations and 0.03- 0.053% in the United States (US)<sup>13; 14</sup>. The apparent variation may in part be a consequence of the variable methodologies employed in the respective studies. Important factors may include the diagnostic criteria, diagnostic accuracy (some studies relied on patient self-reporting), methods of data collection (eg. survey, patient health insurance claims, all of which have their own biases) and the particular populations studied. Low presentation and diagnosis rates due to factors such as embarrassment and a lack of clinical knowledge about the condition are further confounding factors. Consistent with the latter two studies, the incidence is 6/ 10 000 in the US and has significantly increased over the last 40 years, particularly in females (4.3/ 100 000 - 9.6/ 100 000). This increase should be interpreted with caution however as numerous factors, including higher presentation rates and more accurate diagnosis, may have influenced this statistic<sup>15</sup>. The incidence was highest in females aged 20 - 29 years and rapidly diminished in individuals over the age of 50, particularly in post-menopausal women<sup>11; 15</sup>.

#### **1.1.2.1 HS has been associated with obesity and smoking**

Body Mass Index (BMI) appears to confer an impact on disease development and progression<sup>11</sup>. 43 - 77% of HS patients are overweight (BMI 25-30) or clinically obese (BMI > 30)<sup>11; 16-18</sup> and the average disease severity progressively increases in groups of individuals with a BMI of less than 25 (Hidradenitis severity score (HSS) 32), between 25 and 30 (HSS 44) and over 30 (HSS 50), suggesting that weight loss may be of

benefit<sup>16</sup>. Up to 93.7% of patients with HS are current or ex-cigarette smokers compared with 46% of controls<sup>19</sup> (the percentage of HS patients with a smoking history varies from 69% - 93.7% across studies<sup>11; 19</sup>). The average disease severity is highest in current smokers and reduces through ex-smokers and non-smokers suggesting that smoking cessation may be of potential benefit in HS<sup>16; 20</sup>.

### **1.1.3 Clinical presentation**

HS is a chronic inflammatory skin condition that presents with comedones (characteristically paired), papules, pustules, nodules, cysts, abscesses, sinus tracts and fistulae in flexural areas<sup>21</sup>. The condition is associated with significant pain, tenderness, burning, stinging, pruritis, chronic discharge (serous, purulent or blood-stained) and a persistent malodour<sup>4</sup>. Longstanding disease can result in fibrosis, dermal contractures, scarring and a consequent reduction in mobility (see Figure 1.1). The disease targets flexural areas, notably the post-auricular region, axillae, submammary region, abdominal fold, groin, perineum, buttocks and medial thighs. Rarer sites include the scalp, face, neck, back and legs. Of 48 patients in one study, 98% were affected in the groin, 94% in the axilla, 79% buttocks and 34% sub-mammary region<sup>22</sup>. Of 66 female patients in a further study, 56% were affected in the groin, 68% in the axilla, 32% buttocks and 36% sub-mammary region<sup>23</sup>. There appears to be a differential distribution in men and women. Women are more likely to present with axillary, sub-mammary and groin involvement whereas men are more commonly affected in the perineal and perianal region and at rarer sites such as ears, face, chest and back<sup>15; 24</sup>. There is significant inter-individual variation in disease course and severity. Of over 300 cases in one study, 68.2% had “mild” disease (as defined by a method of staging called the Hurley score, discussed in a later section), 27.6% “intermediate” disease and 3.9% “severe disease”<sup>9</sup>. On average, men are more severely affected than women<sup>15</sup>.

Complications associated with HS include fistula formation (affecting the urethra, bladder or rectum), lymphoedema and the development of squamous cell carcinoma (SCC). SCC development is rare in HS (5 of 2119 cases retrospectively analysed in one study) and predominantly arises in the context of longstanding perianal disease<sup>25</sup>. Metabolic syndrome is commonly associated with HS and that association appears independent of patient age, disease severity or disease duration<sup>26</sup>. Features of the metabolic syndrome include central obesity, hypo-HDL-cholesterolaemia, hypertriglyceridaemia, hypertension and hyperglycaemia. Consequently, HS patients may be at high risk of developing cardiovascular disease<sup>26</sup>.



**Figure 1.1. The clinical appearance of HS.** A) An abscess in the sub-mammary region. B) Inflammatory nodules, sinus tracts and scarring in the axilla. C) Paired comedones, inflammatory nodules, sinus tracts and scarring in the axilla. D) Widespread fistula formation and scarring over the buttocks as a result of longstanding, persistent and severe disease. Photos courtesy of Guy's and St. Thomas' National Health Service (NHS) Foundation Trust, all individuals consented to publication.

#### 1.1.4 Histopathology

Common histopathological features include follicular hyperkeratosis, follicular hyperplasia and follicular occlusion with an associated spongiform infundibulofolliculitis<sup>7</sup>. These changes may be associated with follicular dilatation, follicular rupture and the formation of keratin containing cysts (lined by stratified squamous epithelium), abscesses, sinus tracts, granulomas, fibrosis and scarring. Other

features include an interfollicular psoriasiform epidermal hyperplasia, sebaceous gland involution and apocrine gland hyperplasia which is often associated with a periglandular inflammatory infiltrate<sup>6; 27; 28</sup>. Subepidermal, dermal and subcutaneous inflammation is observed in 82%, 76% and 31% of cases respectively and the inflammatory infiltrate comprises predominantly lymphocytes in the early stages followed by neutrophils, multinucleated foreign body giant cells and granulomas as the disease progresses<sup>29</sup>.

### **1.1.5 Phenotypic classification**

The significant inter-individual variability regarding the site of disease, the type of lesions present and the associated systemic symptoms (such as fever, general malaise and arthralgia) during disease flares indicates that there may be multiple, as yet poorly defined, phenotypes within HS. No phenotypic classification existed upon commencement of this project. A recent French study performed a latent class analysis of clinical variables in 618 HS patients which identified three broad phenotypic groups<sup>30</sup>. The first group (axillary-mammary) exhibited mainly breast and axillary involvement associated with hypertrophic scars and was mainly comprised of female patients. The second group (follicular) exhibited ear, axillary, breast, chest, back and leg involvement with a greater tendency to develop comedones, epidermal cysts and pilonidal sinuses. Patients in this group were more commonly obese, male, smokers, had more severe disease and reported a past medical history of acne. The third group (Gluteal) had gluteal disease with superimposed follicular papules and folliculitis. This group exhibited less severe disease, a higher smoking rate but a lower average BMI than the other two groups.

Some rarer but more specific phenotypic sub-groups have been described. Pyogenic arthritis, pyoderma gangrenosum and acne (PAPA) syndrome has been associated with HS and is due to heterozygous mutations in the PSTPIP1 gene<sup>31</sup>. The combination of

PAPA and HS has recently been referred to as PAPASH syndrome<sup>32</sup>. A related phenotypic group comprises individuals with pyoderma gangrenosum, acne and suppurative hidradenitis (PASH syndrome)<sup>33</sup>. SAPHO syndrome can also comprise HS, encompassing a variety of clinical presentations including sterile osteomyelitis, osteitis or arthritis and pustular psoriasis, acne or HS<sup>34</sup>. Two further phenotypic groups have been reported, one with HS and keratitis, the other with HS, spondyloarthropathy and acne conglobate<sup>35; 36</sup>. Keratitis was observed in 4 of 62 HS patients in one case series (HS preceded the development of keratitis by an average of 7 years) and has been reported in several case reports since<sup>37; 38</sup>. The group with HS, acne conglobate and arthropathy develop an axial and/ or large joint arthropathy which is commonly associated with disease flares.

#### **1.1.6 Clinical diagnosis and staging**

The diagnosis is usually based on a clinical history of chronic, painful, inflammatory lesions in characteristic apocrine-gland bearing areas. The first reported diagnostic criterion for HS were proposed in 2000 for the purposes of research<sup>39</sup>. The key principles of those criterion were that individuals required typical and recurrent lesions in characteristic sites. Typical lesions were defined as painful and/ or erythematous papules, nodules, abscesses, dermal contractures or paired comedones. Characteristic sites were defined as the axillae and groin. For diagnosis, individuals required either a) one active lesion and a history of three or more painful or discharging lumps (not specified) in characteristic sites or b) a history of five or more painful or discharging lumps in characteristic sites.

Formal clinical diagnostic criterion were recently proposed at The Second Congress Of The Hidradenitis Suppurativa Foundation mirroring many of those original statements but with minor alterations in the definition of typical lesions and typical sites. They state that individuals require 1) typical lesions (painful nodules, abscesses, sinus tracts,



bridged scars or open comedones) in 2) typical sites (axillae, groin, perineal region, perianal region, infra and inter mammary folds or buttocks) and that 3) the disease must be chronic and recurrent<sup>29</sup>.

Disease severity is most commonly documented using Hurley staging or the Sartorius score. Hurley staging is a comparatively simple system based upon the presence of sinus tracts and scarring. The disease is broadly categorised into three groups, stage 1 (abscesses with no sinus tracts or scarring), stage 2 (multiple abscesses and evidence of sinus tract formation) or stage 3 (multiple abscesses, sinus tracts and scarring affecting a wide area)<sup>40</sup>. The Sartorius score incorporates the number and type of lesions present, distance between lesions, presence of normal skin and the anatomical locations affected<sup>41</sup>. The DLQI score is most frequently used to assess the effect of disease on quality of life<sup>42</sup>.

#### **1.1.7 The impact of the disease on quality of life**

The associated pain, chronic purulent discharge, persistent malodour and the involvement of intimate sites in HS can result in significant patient morbidity. This is often confounded by an initial embarrassment to seek medical advice, delayed or misdiagnosis and a lack of effective treatment options (the average time from first symptoms to formal diagnosis is 12 years)<sup>43</sup>. Hospitalised patients have reported worse quality of life scores than those documented in psoriasis, chronic urticaria and atopic eczema<sup>44</sup>. A survey of 114 patients referred to secondary care revealed an average Dermatology Life Quality Index (DLQI) score of 8.9.<sup>45</sup> Factors such as disease severity, duration and location all impact upon quality of life. Up to 40% of HS patients are diagnosed with depression<sup>15</sup>. The disease can have far reaching social and economic consequences, affecting both relationships and employment. The average patient takes 2.7 days off work a year due to the condition<sup>46</sup>.

### **1.1.8 Treatment of hidradenitis suppurativa**

There is currently a paucity of effective treatment options for HS and very few treatments have been rigorously validated in blinded, randomized and placebo-controlled trials. Mild disease can be managed with topical clindamycin<sup>47</sup>. Intra-lesional steroid injections are of unproven efficacy. More severe disease is more often treated with systemic therapy including antibiotics (tetracyclines, clindamycin, rifampicin), retinoids, dapsone, ciclosporin and oral steroids. That said, oral tetracycline (500mg twice daily) has not been shown to confer any benefit over twice daily topical clindamycin over a three month period, but clinical observations would suggest that it can take up to four months for the oral antibiotics to reach their maximal effect<sup>48</sup>. Clindamycin and rifampicin in combination (both 300mg bd) can significantly improve both disease severity and quality of life<sup>49; 50</sup>. It remains unclear as to whether the clinical benefits of oral antibiotics occur as a result of their anti-bacterial or anti-inflammatory properties. Retinoids appear to be of mixed benefit. Acitretin confers a long-lasting improvement in disease severity in up to 75% of cases whereas isotretinoin is only minimally effective but may alleviate mild forms of the disease<sup>51; 52</sup>. A variety of hormonal therapies (mainly anti-androgen agents) have been used to treat HS in females however their efficacy appears inconsistent and has not been rigorously assessed in clinical trials. Cyproterone acetate, norgestrel (in combination with ethinyloestradiol) and finasteride may confer a benefit in over 50% cases<sup>53-55</sup>. Immunosuppressants including prednisolone, dapsone and ciclosporin can be effective<sup>56; 57</sup>. Colchicine confers a minimal effect and Methotrexate is ineffective<sup>58; 59</sup>.

Anti-TNF agents have recently been trialled in HS. A recent systematic review reported that 50% of individuals on infliximab show a good response, 39% a moderate response and 11% no response<sup>60</sup>. 44%, 35% and 21% showed a good, moderate and no response to adalimumab and 39%, 17% and 44% respectively to etanercept. Four patients have

recently received ustekinumab on which two improved significantly, one moderately and one failed to respond<sup>60</sup>.

Other treatment options include radiotherapy, psoralen and UVA (PUVA) phototherapy, photodynamic therapy, laser therapy (Neodymium:yttrium-aluminium-garnet and carbon-dioxide lasers) and surgery<sup>61-65</sup>. Surgical options include incision and drainage, local excision and extensive excision but recurrence is common. Regrettably, many patients fail to respond to the medical and surgical approaches documented here and more targeted, disease-specific therapies are required to treat this debilitating condition.

### **1.1.9 Conditions associated with hidradenitis suppurativa**

HS can arise as part of a follicular occlusion tetrad comprising HS, acne conglobate, dissecting cellulitis of the scalp and pilonidal cysts/ abscesses<sup>66-68</sup>. 23 - 70% of HS patients report a history of acne vulgaris however there appears to be no correlation between the severity of the two conditions<sup>15; 22; 69</sup>. It has also been associated with conditions including Dowling-Degos disease, Jackson-Lawler pachyonychia congenita, Keratitis-ichthyosis-deafness syndrome, PAPA syndrome, Crohn's disease and Fox-Fordyce Disease. Dowling Degos disease is characterised by macular hyperpigmentation in the flexures (axilla, sub-mammary region, groin) and is caused by heterozygous mutations in Keratin 5 (*KRT5*)<sup>70; 71</sup>. Jackson-Lawler pachyonychia congenita is characterised by pachyonychia, palmar, plantar and follicular hyperkeratosis, steatocystoma multiplex and multiple epidermal cysts over trunk, neck and scalp and is caused by heterozygous mutations in the Keratin 6B and Keratin 17 genes (*KRT6B*, *KRT17*)<sup>72-74</sup>. KID syndrome presents with keratitis, ichthyosis and deafness and is due to heterozygous mutations in Connexin 26 (*GJB2*)<sup>75</sup>. As previously mentioned, PAPA syndrome has been associated with HS and is associated with heterozygous mutations in the *PSTPI1* gene<sup>31</sup>. A *PSTPI1* gene mutation has recently been reported in a patient with PAPA and HS, named PAPASH by the authors<sup>33</sup>. Fox-

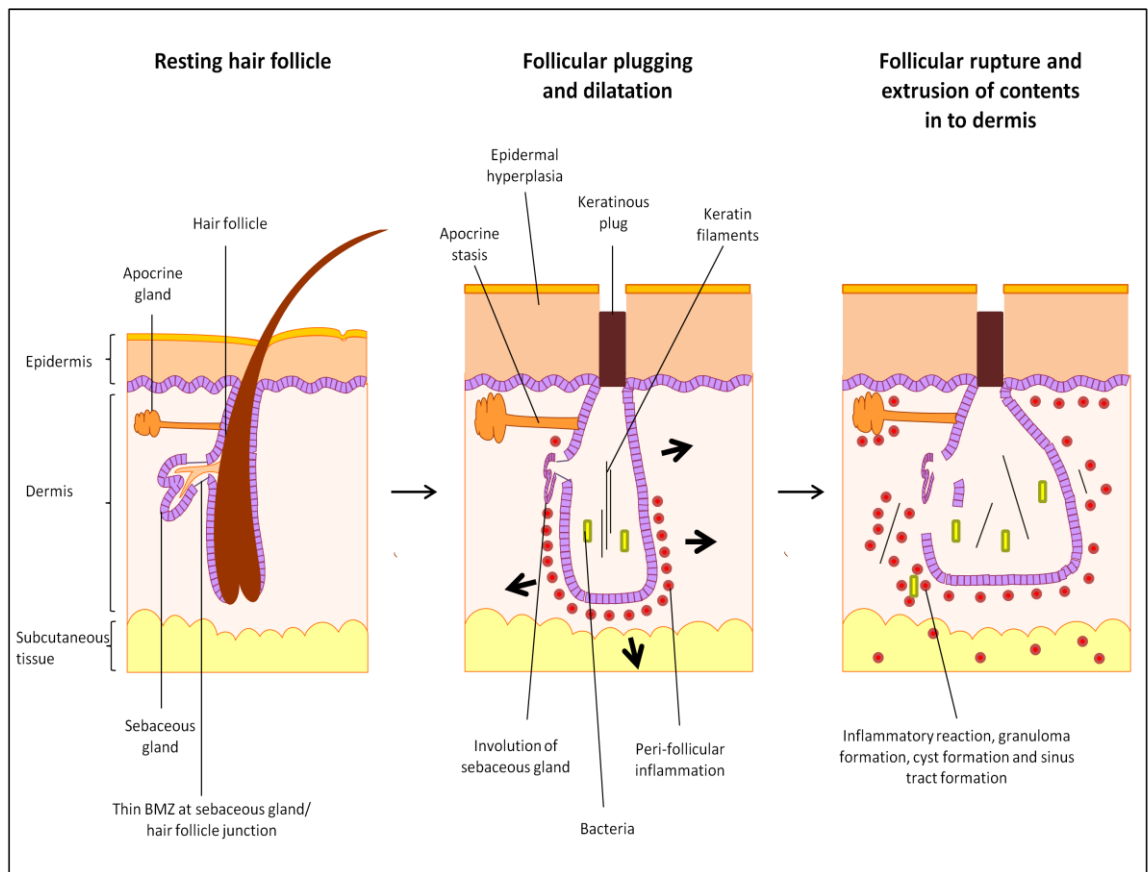
Fordyce disease comprises itchy papules on the chest, axillae and pubic region<sup>76</sup>. The genetic basis of Fox-Fordyce disease is unknown. Crohn's disease is an inflammatory bowel disease that has been associated with variants in multiple genes including pattern-recognition receptors (*NOD2*, *TLR4* and *NLRP3*), genes involved in autophagy (*ATG16L1* and *IRGM*) and genes important in mucosal barrier function such as *DLG5*<sup>77</sup>.

<sup>79</sup>.

## **1.2 Pathogenesis of Hidradenitis Suppurativa**

HS was once considered a disease of the apocrine gland but recent histopathological evidence firmly implicates follicular involvement in disease pathogenesis. Histopathological features observed in over 95% of early lesions (less than three days old) include hyperkeratosis, occlusion of the follicular unit and an associated perifolliculitis. These are often accompanied by dilatation of the hair follicle and subsequent stasis in both the apocrine and eccrine glands<sup>7</sup>. Whilst apocrine stasis, hyperplasia and a peri-glandular inflammatory infiltrate is observed in and around the apocrine glands in 33-90% of HS cases it is universally accompanied by extensive inflammation of the hair follicle, eccrine gland and cystic structures<sup>6; 7; 27</sup>. This implies that changes observed in and around the apocrine glands are likely to arise secondary to primary follicular occlusion<sup>7</sup>. It has recently been observed that there is a thinning of the basement membrane zone (BMZ) around the sebo-follicular junction (neck of the ductus seboglandularis) in HS skin. This area of weakness may predispose to spillage of follicular contents into the dermis and a subsequent inflammatory response<sup>80</sup>. The identification of free keratin filaments in the dermis supports that mechanism and is consistent with follicular rupture playing a key role in disease pathogenesis<sup>81</sup>.

Taken together, these studies suggest that follicular plugging, dilatation and rupture underlies the development of HS (Figure 1.2). The driving forces behind these characteristic histological changes remain unclear however a number of endogenous and exogenous factors have been implicated in the development and propagation of the condition. This section aims to summarise the involvement of the immune system, hormones, bacteria and, of particular relevance to this thesis, genetic predisposition.



**Figure 1.2. The current histopathological model of HS.** A resting pilo-sebaceous unit is represented on the left comprising a thinned BMZ at the sebo-follicular junction. In HS, epidermal and follicular hyperplasia is associated with the development of a keratinous plug in the follicular outlet. This results in the build up of substances including keratin and bacteria within the follicle and consequent dilatation of the pilosebaceous unit (represented in centre). The follicle then ruptures, potentially at the region of weakness at the sebo-follicular junction, extruding its contents in to the dermis. This initiates a significant inflammatory response.

### 1.2.1 Aberrant immune responses in HS

The intense inflammation observed in HS has led many to speculate that aberrant immune responses may play a role in disease pathogenesis<sup>82</sup>. This is supported by the efficacy of immunomodulatory agents such as oral steroids, ciclosporin and more recently anti-TNF $\alpha$  therapy in treating HS<sup>83</sup>. The latter (Infliximab and Adalimumab) specifically imply the involvement of the cytokine TNF $\alpha$  in disease pathogenesis and indeed, TNF- $\alpha$  levels are elevated in the skin and circulation of HS patients<sup>84; 85</sup> and TNF- $\alpha$  receptor 1 and TNF- $\alpha$  receptor 2 expression is increased in affected skin<sup>82</sup>. TNF-

$\alpha$  is capable of recruiting and activating lymphocytes and neutrophils, both of which are prominent in early HS infiltrates, and stimulating the production of a wide array of other pro-inflammatory cytokines<sup>86</sup>. One such cytokine, IL-1 $\beta$  is over-expressed in the skin and circulation of HS patients and levels appear to correlate with disease severity<sup>82</sup>. IL-1 $\beta$  is produced by monocytes and macrophages and is activated by the inflammasome. Inflammasomes are innate immune sensors situated within the cytosol which are triggered by a variety of danger signals including pathogen associated molecular patterns (PAMPS) and damage associated molecular pattern molecules (DAMPS). In the context of HS, these signals may originate from of bacterial sequences, free keratin or DAMPS produced as a result of tissue destruction and scarring. Macrophages and dendritic cells initially recognise these danger signals through pattern recognition receptors situated within the cell membrane such as Toll like receptor 2 (TLR2). Macrophages and dendritic cells harvested from affected HS skin have correspondingly been shown to over-express TLR2<sup>87</sup>. IL-10 expression is also elevated in HS skin, however in contrast to IL-1 $\beta$ , IL-10 is an anti-inflammatory cytokine capable of limiting the response to bacteria and thus reducing overall tissue damage<sup>82</sup>. It can be produced by both cells of the innate and adaptive immune system including macrophages which are prominent in HS infiltrates<sup>82</sup>.

The above evidence suggests that innate mechanisms may initiate the inflammatory response observed in HS. However, cytokine profiles are not consistent across studies and that may in part be due to the exact skin compartments studied and the methods by which expression levels were analysed. A study that focussed explicitly on the expression of cytokines in the epidermis corroborated the finding that IL-10 expression is raised. It also found that the expression of TNF $\alpha$ , TLR-2,3,7,9,  $\beta$ -defensin 2,4, ICAM-1, IL-6, IGF-1,  $\alpha$ -MSH and TGF- $\beta$  was reduced in affected and unaffected epidermis versus control epidermis and epidermis from subjects with other

inflammatory dermatoses including psoriasis and acne vulgaris. The apparent down regulation of these innate markers led the authors to postulate that HS may be a disease of deficient innate immunity<sup>88</sup>. In support of this, monocyte responses to bacteria are reduced in HS and the number of circulating Natural Killer (NK) cells diminishes over time<sup>89; 90</sup>. Furthermore, IL-20 and IL-22 receptor expression is reduced in HS skin and expression of the natural antagonist to those receptors, IL-22 binding protein, is increased<sup>91</sup>.

The involvement of the innate immune system in disease pathogenesis is further supported by the association of HS with Crohn's disease (17% of Crohn's patients have HS)<sup>92</sup>. Like HS, Crohn's disease was once thought of as a direct consequence of bacterial infection. Recent genetic advances however suggest that Crohns disease arises as a result of subtle defects in mucosal immunity. *NOD2* was the first susceptibility gene to be identified which encodes an innate intracellular pattern recognition receptor<sup>79</sup>. The NOD2 receptor recognises bacterial peptides and in turn, enhances intestinal epithelial cell barrier function and mediates chemokine production<sup>93</sup>. In animal models the loss of function variants observed in Crohn's disease disable that receptor function resulting in a defective barrier, bacterial invasion and a chronic inflammatory response (colitis)<sup>94</sup>. Further variants have now been identified in more pattern-recognition receptors (*TLR4* and *NLRP3*), genes involved in autophagy (*ATG16L1* and *IRGM*) and genes important in mucosal barrier function such as *DLG5*<sup>77; 78</sup>.

Antimicrobial peptides represent another mechanism of innate immunity in the skin. They confer anti-microbial, immunomodulatory and pro-inflammatory properties and can promote keratinocyte proliferation and differentiation. Examples include the  $\beta$ -defensins, cathelicidin, psoriasin, ribonuclease 7 and dermcidin<sup>95-98</sup>. The role of antimicrobial peptides in HS is a subject of ongoing investigation. Human  $\beta$ -defensin 2



and 3, cathelicidin (LL-37) and psoriasin have been reported to be over-expressed in HS skin (LL-37 appears particularly upregulated in apocrine epithelium), ribonuclease 7 is reduced and there is no discernible difference in the level of dermcidin<sup>99-102</sup>.

These data provide an insight in to the immune response underlying HS. Epidermal studies and the association with Crohn's disease potentially implicate aberrant innate immunity as a causative factor but the consequential effects on the adaptive response and the exact mechanisms by which any such immune deficiency could result in the severe inflammatory responses observed are yet to be elicited. It remains unclear as to whether immune dysregulation is a primary event, somehow resulting in the characteristic follicular occlusion noted on histology, or in some way a consequence of the cascade of events that follow initial follicular plugging.

Many subtle mechanisms by which the hair follicle contributes to cutaneous immunity and inflammation have been characterised and may be of direct relevance in the pathogenesis of HS. The hair follicle is considered "immune privileged" as, at certain times in the hair cycle, it can protect itself from immune recognition and inflammatory responses via a loss of Major Histocompatibility Complex (MHC) class I and II expression and by expressing CD200, alpha-melanocyte stimulating hormone, transforming growth factor-beta2, macrophage migration inhibitory factor and indoleamine-2,3,-dioxygenase<sup>103; 104</sup>. Multiple immune cells, including Langerhans cells, T cells (CD4 +ve and CD8 +ve), macrophages and mast cells are commonly located in and around the hair follicle<sup>103</sup>. Langerhans cells have been a particular focus of recent study. They form an important bridge between the innate and adaptive immune response and are integral in stimulating cells such as T<sub>H</sub>17 and T<sub>H</sub>2 lymphocytes to react against fungal and bacterial infections<sup>105; 106</sup>. Keratinocyte subsets within the hair follicle have been reported to play an integral role in facilitating the migration of Langerhan cell precursors in to the epidermis. Specific expression of the chemokine

ligands CCL2 and CCL20 appears pro-migratory and CCL8 inhibitory<sup>107</sup>. This recent advance has lead some to speculate that the hair follicle acts as the “gate-keeper” to the epidermis<sup>108</sup>. The emerging immune function of the hair follicle, or indeed any malfunction thereof, may be of direct relevance in HS and this requires investigation.

### **1.2.2 Hormonal influence**

The female predominance, post-pubertal onset, pre-menstrual flares (57% of affected women)<sup>17</sup> and clinical improvement often observed during pregnancy and post-menopause imply a role for hormones in HS pathogenesis<sup>109</sup>. Given the role of androgens in acne vulgaris it has been proposed that they may also play a role in HS. In concordance with this, female patients often suffer with pre-menstrual flares (when progesterone levels are high) and are more likely to present with other androgen related sequelae such as acne, hirsutism and irregular menses (data was never replicated in larger cohorts)<sup>110; 111</sup>. Case reports lend further evidence to the involvement of androgens in HS, documenting that progestogens (androgens) within oral contraceptive agents may induce disease flares. Furthermore, *In vivo* mouse models indicate that androgens may induce a pro-inflammatory state by increasing TLR mediated monocyte expression of TNF $\alpha$ <sup>112; 113</sup>. Therapeutically however, anti-androgens (cyproterone acetate) combined with oestrogens were no more beneficial than oestrogens alone<sup>110</sup> and plasma testosterone, dehydroepiandrosterone and apocrine gland androgen converting enzyme levels are no different in HS patients<sup>23; 114</sup>. Furthermore, the female predominance and the fact that the disease often starts many years after puberty are not consistent with androgen involvement in HS. Overall, the impact of androgens remains uncertain and the mechanism of hormonal involvement in HS requires further clarification.

### **1.2.3 Bacterial infection**

The role of bacteria in HS is controversial. Studies using superficial sampling techniques have shown that a wide variety of bacteria overly HS lesions, including aerobes (*Staphylococcus aureus*, *Streptococcus viridans*), corynebacteria and anaerobes<sup>115</sup>. The reliability of these results is however questionable given that many bacteria on the skin surface are likely to be colonisers rather than directly involved in the disease process. Furthermore, there was very little consistency in the bacterial strains identified across the different studies. One study involving deep needle aspiration of lesions, thus excluding superficial colonising bacteria, revealed no growth in 51% of cases. In the remaining 49% of cases, *Staphylococcus aureus*, coagulase negative staphylococci and occasional anaerobes were detected<sup>116</sup>. Two further studies employed CO<sub>2</sub> laser to dissect through deeper tissue planes. One study focussed on patients with an acute exacerbation<sup>117</sup>, whereas the other looked at patients with chronic disease<sup>118</sup>. Coagulase negative staphylococci were found to be the most abundant organism in superficial and deep planes in both acute and chronic lesions. *Corynebacterium* and anaerobes were the next most abundant organisms in acute lesions (*staphylococcus aureus* not present at any level in the 10 individuals studied) whereas *Staphylococcus aureus* and anaerobes were the next most abundant organisms in chronic disease. Polymicrobial growth was noted in at least one tissue plane in all patients studied and deep tissue cultures were positive in 22 of 25 patients<sup>118</sup>.

Taken together it would appear that bacteria are found within the majority of, but not all, HS lesions. There was considerable inter-individual variation in the number and type of organisms present, although coagulase-negative staphylococci were most common. The lack of consistent organism, the failure of patients to develop infectious complications such as cellulitis, the normal appearance of regional lymph nodes and the

efficacy of oral steroids in HS support a secondary rather than primary role for bacteria in disease pathogenesis<sup>119</sup>. One mechanism by which they may confer such a role is by, upon follicular rupture, being recognised by pattern recognition receptors and triggering an immune response. Eradication of these bacteria may therefore be one mechanism by which antibiotics alleviate symptoms in HS. That said, it remains a matter of debate as to whether antibiotics act via their anti-bacterial actions or through independent anti-inflammatory mechanisms in this context.

#### **1.2.4 Obesity and smoking**

Obesity may have an impact on HS by mechanically increasing friction at flexural sites (thus potentially damaging follicular outlets), increasing sweat retention or increasing the circulating level of pro-inflammatory cytokines (eg. IL-1 $\beta$  and TNF- $\alpha$  are both secreted by macrophages within visceral fat)<sup>10</sup>. Circulating mononuclear cells reportedly exhibit a pro-inflammatory state in obese individuals (IL-6 and TNF- $\alpha$  gene expression is upregulated)<sup>120</sup>. Furthermore, elevated circulatory levels of palmitate in obese individuals may also predispose to inflammation by interacting with TLR's and activating innate immune mechanisms<sup>121</sup>. Alterations to the microbiome in regions of exposed skin may also predispose to disease development<sup>122</sup>.

The exact mechanisms by which smoking contributes to disease pathogenesis are unclear however nicotine has been shown to induce epidermal hyperplasia and follicular plugging<sup>123</sup>. Further mechanisms by which smoking may confer an impact on the disease are that it appears to cause an initial overstimulation and eventual reduction of glandular secretion<sup>124</sup>, results in the secretion of noxious metabolites in sweat, stimulates neutrophil chemotaxis<sup>115</sup> and increases the expression of pro-inflammatory cytokines such as IL- $\beta$  and TNF- $\alpha$ <sup>125</sup>. Furthermore smoking has been shown to reduce cutaneous blood flow, reduce tissue oxygenation and impair wound healing, all of which may be detrimental to the healing of HS lesions<sup>126; 127</sup>. Reactive oxygen species

within tobacco smoke confer immunomodulatory effects by detrimentally affecting phagocytosis, potentially facilitating any infective involvement in disease pathogenesis<sup>128</sup>. Affects on fibroblast migration and activation<sup>129; 130</sup> and an upregulation of MMP1 and MMP3 activity may have a significant impact on the supporting structure in the skin, such as collagen, elastin and proteoglycan<sup>131</sup>, and consequently impact upon the histopathological evolution of the disease.

### **1.2.5 The genetics of hidradenitis suppurativa**

Genetic factors appear to play an important role in disease pathogenesis given that up to 42% of patients report a family history of the condition<sup>22</sup>. Familial studies performed in the 1980's indicated that the condition may be inherited in an autosomal dominant manner<sup>132</sup> and that was confirmed by a follow up study (analysing the same kindreds) 15 years later<sup>39</sup>. The transmission rate in those pedigrees was less than the 50% that you would expect with strict autosomal dominant inheritance and the authors proposed that this may have been as a consequence of the diagnostic criteria used or incomplete penetrance<sup>39</sup>. Variability in disease penetrance may also partly reflect the many intrinsic and environmental factors that can impact upon disease development or progression (discussed above). HS has not been associated with any specific human leukocyte antigen (HLA) types<sup>133</sup>.

### **1.2.6 Mutations in the gamma-secretase genes**

Until recently, only one putative genetic locus had been reported in HS (80cM locus on chromosome 1, 1p21.1-1q25.3)<sup>134</sup>. In 2010, heterozygous mutations were reported in the gamma-secretase genes *PSENEN*, *PSEN1* and *NCSTN* in six Chinese multiplex kindreds, all showing full co-segregation<sup>135</sup>. One of these genes, *NCSTN*, lies within the previously reported region of linkage on chromosome 1<sup>134</sup>. Mutations have since been reported in two British, six Chinese, one Japanese and three French multiplex kindreds as well as four apparently sporadic cases<sup>135-142</sup>. To date, eighteen mutations have been

reported in *NCSTN*, three in *PSENEN* and one in *PSEN1*<sup>135-141</sup> of which four are nonsense mutations, seven result in frameshifts, seven in altered splicing and four are missense mutations. The data presented in this thesis contributed to this genetic summary and, to our knowledge, represents the most comprehensive functional mutational analysis performed to date.

### 1.3 Gamma-secretase

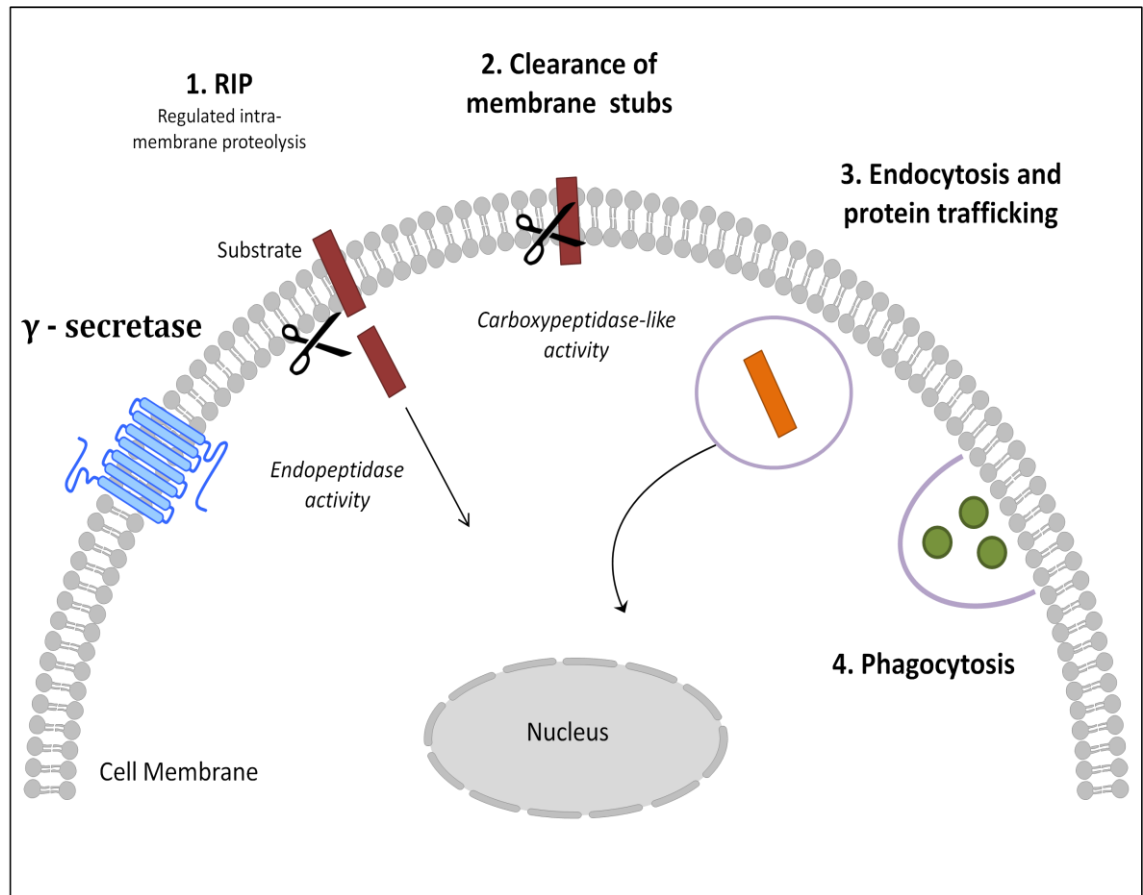
Gamma-secretase is an aspartyl intra-membrane protease complex capable of hydrolysing and cleaving in excess of 50 type-1 transmembrane proteins including amyloid-precursor protein (APP), Notch receptors, N-cadherin, E-cadherin, neuregulin, Erb4 and Nectin-1 $\alpha$ <sup>143-145</sup> (full table of substrates shown in Appendix 1). These membrane proteins have large ectodomains which are cleaved by a variety of sheddases. The remaining intramembranous portions are then cleaved by gamma-secretase (endopeptidase activity,  $\epsilon$ -cleavage), releasing the intracellular component of each protein. Some of these intracellular fragments, such as Notch intracellular domain (NICD), play important biological roles within the cell<sup>146</sup>. Following the initial  $\epsilon$ -cleavage, some substrates undergo further processing by gamma-secretase whereby the remaining intramembranous stub is sequentially cleaved to form by-products of varying sizes (carboxypeptidase-like activity). This is evident in both APP and Notch processing.

Despite the intricate and complex mechanisms by which this enzyme interacts with and cleaves its respective substrates it appears to demonstrate a remarkable lack of substrate specificity. Correspondingly, no common recognition sequence is shared between all substrates. Four factors appear to determine whether a membrane protein will be targeted by gamma-secretase, orientation (type I), physical conformation (must contain a hydrophobic  $\alpha$ -helix), prior ectodomain shedding (gamma-secretase will not interact until this cleavage has occurred) and the residual length of the ectodomain following initial cleavage by another sheddase (proteins with ectodomains of less than 50 amino acids are effectively cleaved)<sup>147</sup>. Furthermore, substrates are required to have “permissive” transmembrane and cytoplasmic domains that may, for instance, enable any conformational changes that are required for proteolysis<sup>148</sup>.

Gamma-secretase has traditionally been viewed as playing two key roles. The first is to catalyse cleavage of the potentially biologically relevant intracellular portions of substrates (eg. NICD in the context of Notch signalling), a process known as regulated intra-membrane proteolysis (RIP)<sup>149</sup>. The second is to remove intramembranous proteins from the cell membrane and thus maintain healthy cell cycling and homeostasis<sup>150</sup>. The biological role of short peptides produced as part of that process (via carboxypeptidase-like processing of the remaining intramembranous stubs) remains unclear (eg.  $\beta$ -amyloid (A $\beta$ ) in the case of APP processing).

The components of gamma-secretase are conserved through metazoans, higher plants and even slime mould. Work in plants and mould indicates that the enzyme may have functions outside of membrane clearance and RIP that are integral to normal cell function. No enzyme substrates have been identified in plants, moulds or mosses yet null mutations in presenilin (a component of gamma-secretase) affect growth patterns, chloroplast movement, membrane internalisation and endosome trafficking in moss (*Physcomitrella patens*)<sup>151</sup> and affect phagocytosis in slime mould<sup>152</sup>. Similar phenotypes were observed in nicastrin (another component of gamma-secretase) null mutants. In addition to RIP and removing and recycling cell membrane proteins, gamma-secretase would therefore appear to be involved in a diverse range of cellular mechanisms including endocytosis (including receptor mediated endocytosis), protein trafficking (eg. APP) and phagocytosis (further confirmed by looking at macrophage function in mice)<sup>153-156</sup>. The functions of RIP and re-cycling cell membrane proteins may only have developed as more complex membrane proteins evolved<sup>152; 157</sup>. The cellular functions of gamma-secretase are shown in Figure 1.3.



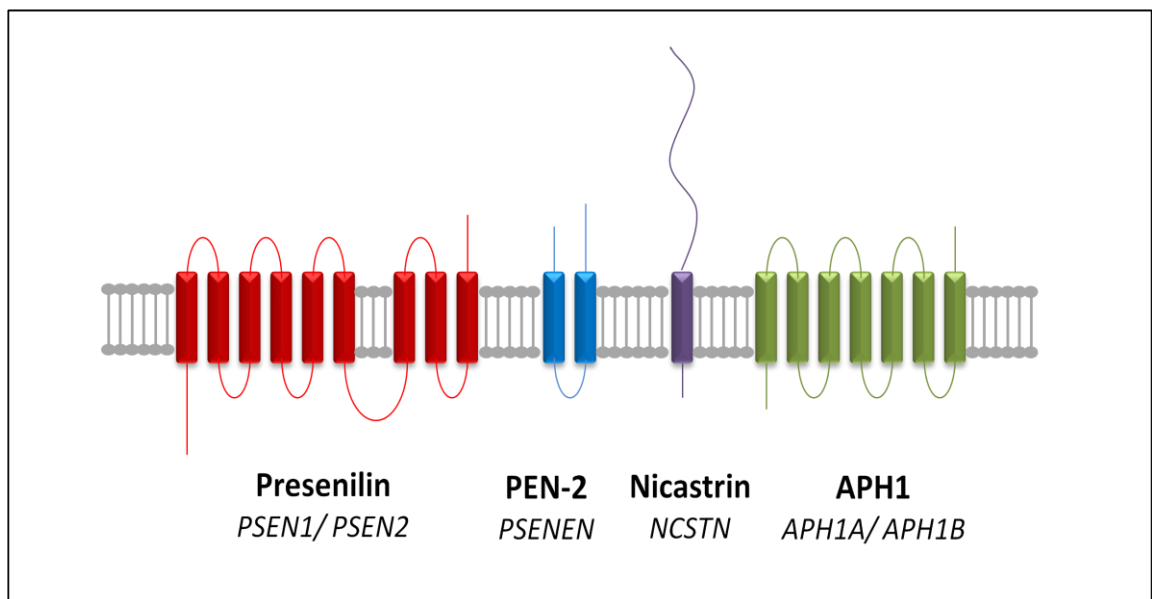


**Figure 1.3. The key roles of gamma-secretase in cellular functioning.** Gamma-secretase is involved in regulated intramembraneous proteolysis, the clearance of protein membrane stubs from the cell membrane, endocytosis, protein trafficking and phagocytosis.

### 1.3.1 The composition of the gamma-secretase complex

Gamma-secretase is a protein complex composed of at least four fundamental sub-units comprising presenilin, nicastrin, PEN-2 and APH-1. The combination of these four proteins in a 1:1:1:1 ratio forms an active enzyme complex<sup>158; 159</sup> (Figure 1.4). These proteins are encoded by the genes *PSEN1* or *PSEN2*, *NCSTN*, *PSENEN* and *APH1A* or *APH1B* respectively. Because presenilin can be encoded by *PSEN1* or *PSEN2* and APH-1 can be encoded by *APH1A* or *APH1B* there are at least four different protein combinations that can potentially form an active enzyme complex and this is further compounded by alternative splicing of *APH1A* (*APH-1aS*, *APH-1aL*). Immunoprecipitation experiments have indeed confirmed the existence of these

alternative complexes (weight varies from 380 – 550kDa) and furthermore, that they can co-exist in the same cell type.<sup>160</sup> The biological rationale for the existence of multiple subtly different complexes is as yet unknown, however it maybe that each confers a degree of substrate specificity<sup>161</sup>. Interestingly PSEN1, PSEN2 and APH-1B demonstrate a degree of tissue specificity (eg. PSEN1 expression higher in testis, PSEN2 expression higher in liver, APH-1B expression higher in brain) whereas APH-1A appears uniformly expressed (both splice variants)<sup>160</sup>.



**Figure 1.4. The gamma-secretase complex.** The gamma-secretase complex is an intramembranous multi-protein complex composed of 4 subunits, presenilin, PEN-2, nicastrin and APH-1, encoded by *PSEN1/ PSEN2*, *PSENEN*, *NCSTN* and *APH1A/ APH1B* respectively.

The production, maturation, migration, stability and eventual activity of these individual proteins appears to be closely intertwined such that any imbalance or dysfunction of any one unit will impact on the others and ultimately affect enzyme activity<sup>161</sup>. Most studies demonstrate that if one subunit is missing (through knockdown or knock-out) then the other proteins are proteasomally degraded, thus abolishing enzyme activity<sup>162</sup>. Down regulation of any one component inhibits enzyme activity<sup>163-166</sup> and over-expression of all four increases enzyme activity<sup>145; 164; 165</sup>.

Immunoprecipitation of active enzyme complexes confirms that one of each protein forms an active complex<sup>145</sup>. That said, all four subunits are not essential for enzyme activity under certain experimental conditions. Mutagenesis of specific residues within PSEN1 (S438P + one of 15 other mutations) enables minimal gamma-secretase activity in the absence of nicastrin *in vitro* (in yeast and mouse fibroblasts)<sup>167</sup>. The S438P mutation in PSEN1 trans-membrane (TM) segment 9 was proposed to induce a conformational change in the complex that was sufficient to allow substrates to access to the catalytic site in the absence of NCSTN. It was noteworthy however that the mutant presenilin fragments were unstable in the absence of nicastrin and that enzyme activity was notably reduced compared to wild-type complexes containing nicastrin. A further study demonstrated that mutant PSEN1 (again lacking part of the TM9 sequence) alone and wild-type presenilin and PEN-2 in combination can exhibit some enzyme activity when re-constituted in proteoliposomes<sup>168</sup>. In summary, mutant forms of presenilin, alone or in combination with PEN-2/ APH-1, but not NCSTN, are still catalytically active and wild-type presenilin/ PEN-2 complexes can confer some enzyme activity. It should be noted however these experiments were performed under very specific experimental conditions that do not necessarily reflect the true *in vivo* environment of gamma-secretase and mainly studied mutant proteins. Even the work with wild type presenilin does not reflect the complicated assembly, maturation, trafficking and general stability of gamma-secretase in the body. Whilst it is therefore possible that not all wild type complex components are necessary to confer strict catalytic activity under experimental conditions they appear essential in producing and maintaining stable, mature and active complexes *in vivo*. Correspondingly, mRNA expression of each complex component very tightly mirrors that of the other three across different tissues. The mechanisms underlying this coordinated transcriptional regulation are poorly understood<sup>160</sup>.

Presenilin was the first component to be characterised following the identification of *PSEN1* and *PSEN2* mutations in familial Alzheimer's disease (discussed further in section 1.3.7)<sup>169</sup>. Despite exhibiting the features and sequence of an aspartyl protease, presenilin alone exhibited no catalytic activity<sup>170</sup>. Consequently it was hypothesised that further proteins may be required to facilitate enzyme activity. Co-immunoprecipitation of presenilin lead to the subsequent identification and purification of nicastrin and a screen for proteins that are integral to Notch-pathway signalling in *C.elegans* identified the final two components, PEN-2 and APH-1<sup>163; 171; 172</sup>. The exact function of each of the individual components within the gamma-secretase complex is still poorly understood but is discussed in the following sections.

#### **1.3.1.1 Presenilin**

Presenilin (encoded by *PSEN1* or *PSEN2*) is a multi-pass membrane protein comprising nine trans-membrane (TM) domains which forms the catalytic core of the gamma-secretase complex. The immature protein is produced in the endoplasmic reticulum (ER) which is subsequently cleaved (endoproteolytic cleavage between PSEN1 Thr291 and Ala299) to form a C-terminal fragment and an N-terminal fragment, referred to as mature presenilin<sup>173; 174</sup>. This cleavage is associated with a conformational change of the complex which facilitates enzyme activity. Each fragment harbours one of two key and highly conserved aspartate residues (TM 6 and 7, aa's 256 and 385) which together form the active catalytic unit<sup>175</sup>.

One of these aspartate residues lies within a highly conserved Glycine-X-Glycine-Aspartic acid motif (aa 382-385, where amino acid X varies between species) and whilst the aspartic acid residue within this motif is critical for catalytic activity the two glycine residues and the variable X residue also confer important and diverse roles in enzyme activity. Mutational analysis of these residues suggests that the first (PSEN Gly382) is required for effective endopeptidase activity and the second (PSEN Gly384) for

effective carboxypeptidase-like activity<sup>143</sup>. They may also play a role in substrate specificity<sup>176</sup>. The polarity, charge and size of the variable “X” residue effects carboxypeptidase-like activity, possibly by altering the conformation of the protein<sup>177</sup>. Met139 is a further specific amino acid which seems to be selectively important in carboxypeptidase-like but not endopeptidase activity inferring that it may be in close proximity to Glycine-X-Glycine-Aspartic acid motif in the three dimensional structure of presenilin<sup>178</sup>.

A Pro-Ala-Leu sequence in TM9 (the PAL motif) is another highly conserved and functionally important region of presenilin. Mutagenesis of this region prevents nicastrin binding and renders the enzyme inactive<sup>179; 180</sup>. Cross-linking experiments and more recently x-ray crystallography have shown that this motif lies very close to the catalytically critical aspartate residue in TM6 (Asp256)<sup>181; 182</sup>. TM9 may therefore be important in facilitating substrate migration to the active catalytic site within presenilin<sup>183</sup>. The fact that presenilin can be co-precipitated with substrates infers that the two are in prolonged contact<sup>184</sup>. This potentially supports the hypothesis that a conformational change of the complex is required to move the substrate to the cleavage site and may imply that there is a significant distance between any potential substrate binding site and the catalytic cleavage site<sup>185; 186</sup>.

Presenilin appears to form the main cohesive link in the gamma-secretase complex as it is the only subunit that can interact with all of the other constituents. PSEN1 can interact with NCSTN and PEN-2 in the absence of APH-1A, NCSTN and APH-1A in the absence of PEN-2 and with APH-1A and PEN-2 in the absence of nicastrin. NCSTN and APH-1A can interact with one another in the absence of PSEN1 but not with PEN-2<sup>162</sup>. Presenilin is therefore required to mediate any reaction between APH-1A/NCSTN and PEN-2. The PSEN1-NCSTN interaction is particularly interesting as PSEN1 needs

NCSTN for stability and NCSTN requires PSEN1 to exit the ER and mature (mature nicastrin is not produced in PSEN1 deficient mouse embryonic fibroblasts (MEF))<sup>99; 112</sup>.

The function of presenilin within the complex is further complicated by the existence of two isoforms, PSEN1 and PSEN2. All mature gamma-secretase complexes contain either PSEN1 or PSEN2 but not both<sup>160</sup>. These isoforms share 67% of their sequence but PSEN1 appears to be the more abundantly expressed and dominant isoform<sup>187 105; 106</sup>. *In vivo*, *PSEN1* knockout is lethal whereas *PSEN2* knockout mice demonstrate a mild phenotype with apparent lung fibrosis<sup>188; 189</sup>. *In vitro*, PSEN1 and 2 deficient cells display no enzyme activity (APP as substrate), PSEN1 deficient cells display very little enzyme activity and PSEN2 deficient cells display almost normal activity<sup>190</sup>. This may in part be due to the fact that PSEN1 is required for the effective trafficking of NCSTN out of the ER, thus its absence would impact upon PSEN2 complex formation. That effect does not appear to be reciprocated in PSEN2 deficient cells where normal PSEN1 complexes are assembled<sup>160</sup>. Once formed, there is conflicting data regarding the relative efficacy of each complex in cleaving APP. Purified PSEN1 and PSEN2 gamma-secretase complexes appear to show no significant difference in APP processing but PSEN2 complexes do however seem less likely to form than their PSEN1 counterparts<sup>166; 167</sup>. There is some evidence to suggest that the nature of the assembled isoform differentially facilitates binding with varying complex interactors (eg. proteins that regulate gamma-secretase activity, discussed further in section 1.3.4) and may ultimately direct complex activity towards very separate signalling cascades, for example PSEN2 appears particularly important in PDGF signalling.

Presenilin is clearly integral to gamma-secretase enzyme activity but there is emerging data suggesting that it may also be involved in biological processes unrelated to the complex. Independent functions may include protein trafficking (including proteins

such as APP, telencephalin and the Trk receptor) and a role in regulating intracellular calcium levels<sup>191 98; 108</sup>.

### 1.3.1.2 Nicastrin

Nicastrin (encoded by *NCSTN*) is a single pass trans-membrane protein composed of a large hydrophilic ectodomain (669aa), a short hydrophobic trans-membrane domain (20aa) and a short hydrophilic C-terminus (20aa)<sup>171</sup>. It is thought to be involved in complex assembly, migration, maturation and substrate recognition and accounts for just under half of the molecular weight of the gamma-secretase complex<sup>144; 192</sup>. An endoglycosidase-H-sensitive glycosylated precursor protein (immature form, 110kDa) is synthesised in the ER which undergoes glycosylation in the Golgi apparatus and sialylation in the trans-Golgi network (mature form, 130kDa, stable for up to 24 hours)<sup>186</sup>. The ratio of immature to mature nicastrin varies in different cell types, for example it is high in fibroblasts and low in neurons<sup>186</sup>. Mature nicastrin is preferentially found in gamma-secretase complexes (PSEN1 has a higher affinity for mature nicastrin than immature nicastrin<sup>162</sup>) however immature forms do not detrimentally affect enzyme activity<sup>186</sup>. Nicastrin can interact with PSEN and APH-1 via its transmembrane and ectodomain<sup>193 194</sup>. APH-1 binding is essential in forming the initial nicastrin-APH-1 subcomplex in the ER which precedes full assembly of the complex (section 1.3.2) and PSEN1 binding is required for immature nicastrin to exit the ER<sup>186</sup>.

The importance of nicastrin in overall complex assembly and activity has been demonstrated in animal studies. Fibroblasts harvested from nicastrin knock-out mouse embryos exhibit unstable and largely inactive gamma-secretase complexes (studying APP as a substrate)<sup>195; 196</sup>. Conversely, over-expression of nicastrin is sufficient to enhance gamma-secretase activity *in vitro* and *in vivo* (enhances APP processing)<sup>197</sup>. Over-expression was also shown to induce behavioural changes (suggestive of a lack of memory) and increase the expression of the other complex components in mice *in*

*vivo*<sup>198</sup>. The apparent memory deficit was postulated to occur as a consequence of altered APP processing and subsequent production of the amyloidogenic  $\beta$ -amyloid 42 (A $\beta$ 42) peptide<sup>198</sup>.

The exact functions of nicastrin within the gamma-secretase complex remain poorly defined. A conformational change in the ectodomain upon integration in to the gamma-secretase complex is required for normal enzyme activity<sup>199</sup>. Three domains within the Nicastrin protein are highly conserved between animal species (amino acids 306 – 360, 419 – 458, 625 – 662). The first of these has been a focus of investigation because it contains a very highly conserved and seemingly important DYIGS motif (aa 336-340)<sup>171</sup>. Because this region of the ectodomain resembles aminopeptidase structures it is referred to as the DYIGS and peptidase homologous (DAP) region<sup>171</sup>. Mutagenesis of the DYIGS motif to AAIGS increases gamma-secretase carboxypeptidase-like processing of APP (increasing A $\beta$  peptide secretion)<sup>171</sup>. Conversely, homozygous deletion of residues 312-369 and 312-340 (containing DYIGS) reduces APP processing (significant reduction in A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> levels, the effect was more pronounced with the deletion of residues 312-369). These larger deletions appeared to affect the nicastrin-PSEN1 interaction.

Mutation of the nearby E333 residue completely inactivates the gamma-secretase complex<sup>192</sup>. It was proposed that this residue binds to the amino terminus of membrane proteins (following initial shedding of their ectodomains) and may act as a substrate recognition site for gamma-secretase<sup>192</sup>. A later study disputed this claim, demonstrating that whilst mutations affecting this residue reduce nicastrin glycosylation, incorporation in to the complex and the amount of complex present they did not in fact alter the enzyme activity per complex<sup>200</sup>. Furthermore, deletion of residues 312-340 and 312-369, both containing E333, did not appear to affect substrate binding<sup>171; 199; 201</sup> and mutagenesis of other residues within the aminopeptidase-like



domain of nicastrin that are integral to the catalytic function of related aminopeptidases conferred no significant effect on substrate binding<sup>200</sup>. Despite these data not supporting a substrate-recognition role for this part of the ectodomain it was later shown that NCSTN and APH-1A can both directly interact with APP in the absence of PSEN1 (NCSTN with a higher affinity).<sup>162</sup> Furthermore, PSEN1 seems unable to bind APP in the absence of APH-1 or NCSTN inferring that the NCSTN-APH-1A subcomplex plays a key role in substrate recruitment to the gamma-secretase complex<sup>162</sup>. Furthermore, monoclonal antibodies have recently been developed which target the ectodomain of NCSTN and neutralise gamma-secretase activity *in vivo* and *in vitro*<sup>202; 203</sup> with immunoprecipitation experiments indicating that one of the antibodies inhibits nicastrin-substrate binding<sup>202</sup>. Specific mutagenesis of the L571 residue within the antibody binding region was found to inhibit substrate binding and consequently inhibit enzyme activity without affecting complex assembly<sup>203</sup>. In summary, the NCSTN-APH1 subcomplex and more specifically the L571 residue within the ectodomain of NCSTN appear to play a key role in gamma-secretase substrate recognition. Interestingly, combined C213S and C230S mutations affecting the NCSTN ectodomain affected APP but not Notch cleavage suggesting that different regions of the ectodomain may exhibit substrate specificity<sup>198</sup>. G365 and S425 represent two further functionally important amino acids within the ectodomain, with mutagenesis experiments demonstrating that they are important in complex formation but do not affect enzyme activity upon complexes forming.

Nicastrin appears integral to optimal gamma-secretase activity however, as with presenilin, it may be able to function independently of the complex. In support of this, levels of nicastrin and presenilin are not co-ordinately regulated<sup>204</sup> and some nicastrin and presenilin is not co-localised in cells. There is early evidence to suggest that nicastrin may be involved in p53 mediated cell death<sup>205</sup>

### 1.3.1.3 PEN-2

PEN-2 (encoded by *PSENEN*) is a 101 amino acid membrane protein comprising two extra-cellular termini, two trans-membrane domains, and one cytoplasmic loop. It is the smallest component of the gamma-secretase complex and the last to be assembled. It stabilises the complex (the other components are rapidly degraded in the absence of PEN-2) and may induce a conformational change in the active site of presenilin to facilitate enzyme activity. PEN-2 binds to an “NF” motif in PSEN1 TM4<sup>206</sup>. The length and sequence of the C-terminal domain and the sequence of the proximal two-thirds of TM1 appear integral to that interaction and to PEN-2 providing stability to the complex<sup>207; 208</sup>.

Alterations in PEN-2 expression affect the maturation of presenilin. Inhibition of PEN-2 increases the ratio of full length PSEN1 to PSEN1-CTF/ PSEN1-NTF and PEN-2 knockout results in no expression of PSEN1-CTF/ PSEN1-NTF<sup>164</sup>. On this basis it was proposed that PEN-2 is required for the endoproteolysis and thus maturation, stabilisation and activation of presenilin<sup>207; 208</sup>. However, recent work has shown that PSEN1-CTF and PSEN1-NTF are still produced in the absence of PEN-2 but are rapidly degraded by the proteasome (reversed in a time-dependent manner with proteasome inhibitors). This also appeared to be the case in NCSTN or APH1 knockout cells. Thus it was proposed that PEN-2 (and NCSTN and APH-1) is required to facilitate effective endoproteolysis of presenilin by stabilising the PSEN1-CTF and PSEN1-NTF products, rather than by having any direct involvement in the actual cleavage<sup>162</sup>. As mentioned previously, PEN-2 and presenilin alone are sufficient to support some catalytic activity in proteoliposomes<sup>168</sup>. This suggests that NCSTN and APH-1 are not essential for the stabilisation of PSEN1-CTF and PSEN1-NTF in that experimental setting and that PEN-2 may be the more dominant of the three in fulfilling that role<sup>168</sup>. This is supported by the increased ratio of full length presenilin to PSEN1-

CTF/ PSEN1-NTF observed in PEN-2 knockdown cells. Increased levels of full length PSEN1 suggests that the Nicastrin-APH1A complex alone can stabilise and protect the full length PSEN1 protein but that it is not sufficient to stabilise the products of endoproteolysis without PEN-2<sup>162</sup>.

#### **1.3.1.4 APH-1**

APH-1 (encoded by *APH1A* or *APH1B*) is a multi-pass membrane protein with seven trans-membrane domains. Along with nicastrin, it appears to act as a scaffold upon which the complex can be assembled and may play a role in substrate recognition<sup>209</sup>. NCSTN and APH-1A have a very high affinity for one another and if the expression of one is reduced it has a profound effect on the expression of the other<sup>162</sup>. PSEN1 and PEN-2 appear to confer further stability to these components<sup>162</sup>. Aside from forming an important scaffold, the NCSTN/ APH-1 interaction appears to facilitate effective transfer of nicastrin to the outer cell membrane<sup>172</sup> and, as mentioned previously, may play a role in substrate binding<sup>162</sup>. Two highly conserved histidine residues (H171, H197) in TM5 and TM6 respectively are integral to complex assembly, stability and ultimately activity<sup>210</sup>. Those same residues are integral to substrate recognition and binding<sup>211</sup>.

The differential roles of APH-1A (and associated splice forms APH-1aS and APH-1aL) and APH-1B remain unclear however APH-1A would appear to be the more dominant isoform given that *APH1A* knockout results in lethality which is not seen with *APH1B* knockout mice<sup>212; 213</sup>. APH-1A deficiency reduces the expression of the other enzyme components, increases the expression of APH-1B and inhibits enzyme activity. Conversely, APH-1B deficiency confers no effect on the expression of APH-1A, presenilin, nicastrin or PEN-2 and does not affect enzyme activity. Despite this, there is some evidence to suggest that APH-1B may play a dominant role in APP processing (A $\beta$  production)<sup>212</sup>. It has been postulated that the APH-1 subtype, and therefore

potentially the splice variants of APH-1A, in each complex may affect which regulatory proteins are able to bind to and thus regulate complex activity<sup>214</sup>.

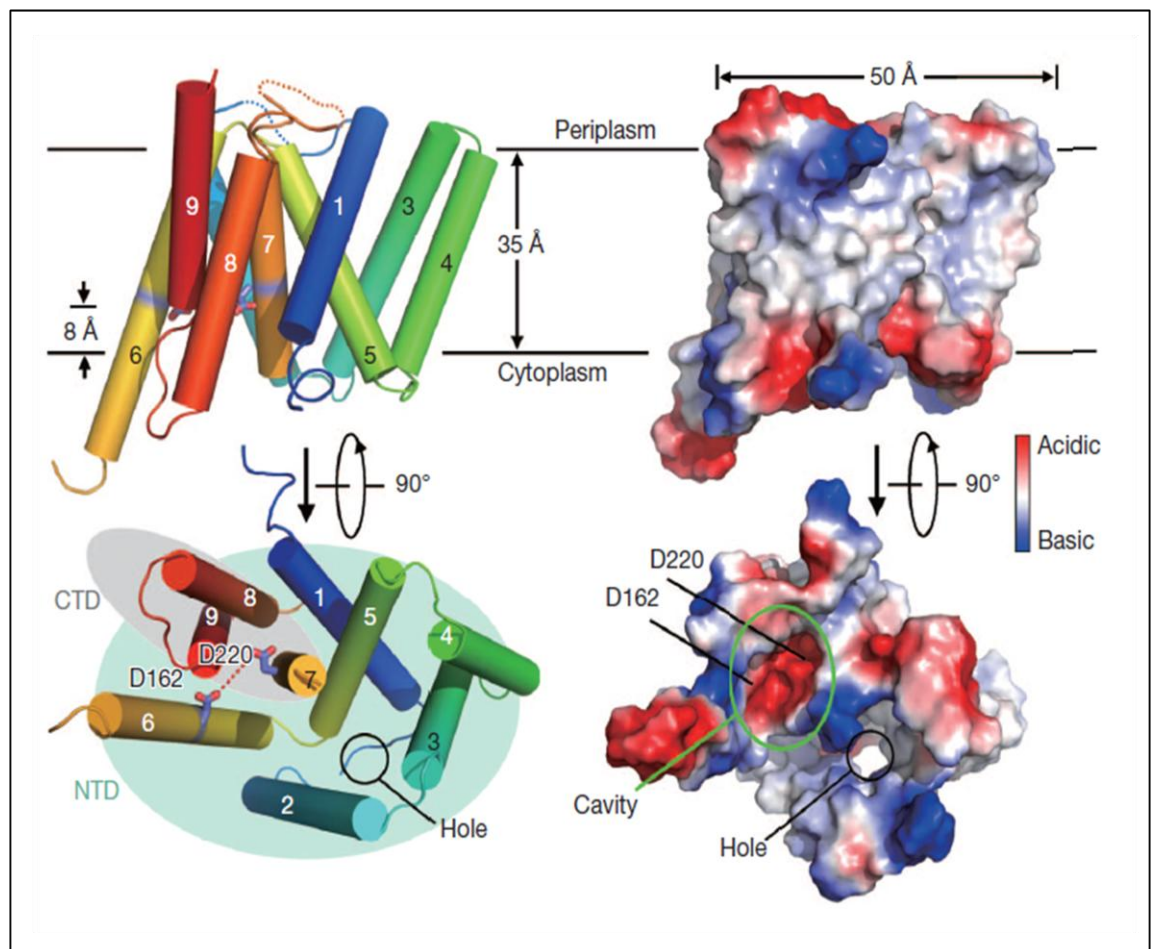
### **1.3.2 The assembly of the gamma-secretase complex**

All of the gamma-secretase components are synthesised in the ER. The formation of a sub-complex comprising immature nicastrin and APH-1 appears to initiate complex assembly<sup>215</sup> however it should be noted that presenilin can form a sub-complex with APH-1 in the absence of nicastrin<sup>162</sup>. Presenilin and PEN-2 are then sequentially added to complete the complex and presenilin is cleaved (auto-proteolysis) to form PSEN1-CTF and PSEN1-NTF. Any components that are not incorporated in to forming complexes are subject to rapid proteosomal degradation. Only a minority of the gamma-secretase complexes (up to 5%) migrate to the Golgi apparatus and mature (nicastrin is subject to further N-glycosylation and matures in the Golgi). These mature complexes are then trafficked to the cell membrane and endosomes where they are catalytically active<sup>189</sup>. Whilst complex constituents are found throughout the ER and Golgi apparatus it is intriguing that enzyme activity is only detectable at or very close to the cell membrane<sup>216; 217</sup>. These active complexes constitute a minority of the total number of complexes within the cell. Up to 95% are seemingly inactive and remain within or cycle between the ER and Golgi<sup>218; 219</sup>. Stringent quality control mechanisms appear to be in place to ensure that only correctly folded and mature complexes can exit the ER. Each component of the complex appears to contain at least one ER retention sequence and it has been proposed that the correct assembly of the complex results in the masking of these specific motifs, thus allowing exit from the ER<sup>179; 220-222</sup>. Further mechanisms are in place to restrict complex formation in the ER. Rer1 protein for example regulates complex formation by competing for binding with NCSTN and PEN-2<sup>221; 222</sup>.

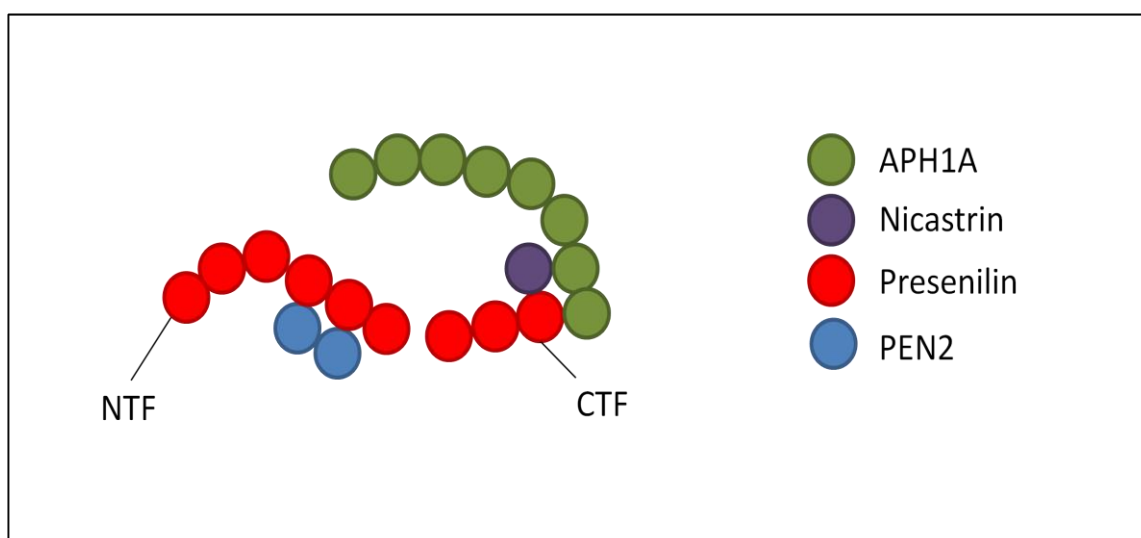
### 1.3.3 The structure of the gamma-secretase complex

Regulated intra-membrane proteolysis describes the process of protein cleavage within the lipid bilayer by a membrane-bound protease. These membrane-bound proteases can be divided into three broad groups, aspartyl proteases (presenilin and signal peptide proteases), metalloprotease site 2 protease and serine protease rhomboid. Whilst it is known to contain 19 trans-membrane domains the three dimensional structure of the gamma-secretase complex has been hard to elicit<sup>144</sup>. Studies using negative stain electron microscopy and electron cryo-microscopy have produced slightly conflicting results regarding the exact shape and volume of the complex. Overall, they indicated that the complex adopts a cylindrical structure harbouring a fully or partially patent aqueous cavity<sup>223; 224</sup> and that the complex weighs around 230kDa, (consistent with the theory that each complex is made up of each of the four component parts in a 1:1:1:1 ratio)<sup>224</sup>. This cylindrical globular structure has recently been corroborated via x-ray crystallography of the archeal presenilin homologue (PSH). Whilst this is not presenilin, the highly conserved amino acids account for just under half of the total aligned sequences and the catalytic motifs are completely conserved. The nine TM domains are arranged at varying angles in the membrane and the NTF arches like a horse-shoe over the CTF fragment. There is a fully patent pore created by TM 2, 3, 5 and 7 (Figure 1.5). The catalytically active residues (Asp162 and Asp220 in TM 6 and 7) lie at the bottom of a separate cytoplasm facing cavity (8Å from lipid membrane surface). They face one another and are situated 6.7Å apart. Water would be available through this cavity which is essential for peptide cleavage. The distance between aspartyl residues in related aspartyl proteases is less than that of a hydrogen-bond (just over 3.1Å) which is required for successful proteolytic activity. The larger distance in the archael presenilin homologue supports the notion that a conformational change is required, perhaps on substrates binding to the gamma-secretase complex, for the presenilin to become active.

The resting conformation may in fact act to prevent non-specific cleavage of membrane proteins. Substrate entry is proposed to be between TM6 and TM9. Via structural modelling these authors predicted that PEN-2 would bind to the opposite side of PSEN-CTF to nicastrin and APH-1 which is in accordance with the aforementioned process of complex formation (Figure 1.6)<sup>182</sup>. Overall, these imaging studies suggest that an aqueous cavity within the protease facilitates intramembranous cleavage of membrane proteins.



**Figure 1.5. The 3D structure of the archael presenilin homologue (PSH).** This figure was taken directly from a paper by Li *et al.*, 2013<sup>182</sup>. The nine trans-membrane domains of PSH are orientated at varying angles. The protein contains a fully patent pore surrounded by TM domains 2,3,5,7. The catalytically active aspartyl residues D220 and D162 face one another across a cytoplasm-facing cavity. Substrates are proposed to enter between TM6 and TM9.



**Figure 1.6. The predicted structure of the gamma-secretase complex.** Mature presenilin is comprised of PSEN-CTF and PSEN-NTF. The C-terminal end of PSEN-CTF is predicted to be in contact with APH-1A and NCSTN and the N-terminal end in contact with PEN-2.

### 1.3.4 Proteins that associate with the gamma-secretase complex

The molecular weight of the complex varies significantly between studies, often far surpassing the cumulative predicted molecular weight of the four constituents. This may in part be due to the array of molecular techniques adopted however it may also be due to a variety of different proteins associating with the complex at the time of measurement. Multiple proteins have now been identified which can affect complex assembly, trafficking, activity and degradation but overall understanding of regulatory mechanisms governing enzyme activity remains in its infancy.

A number of proteins have been shown to support complex activity. These comprise parkin, G-protein-coupled receptor 3,  $\beta$ -arrestin 1,  $\beta$ -arrestin 2, ATP-binding cassette transporter 2 (ABCA2), gamma-secretase activating protein (GSAP) and NADH dehydrogenase iron-sulfur protein 7 (NDUHF7)<sup>225-231</sup>. Parkin interacts with the presenilin promoter to upregulate PSEN1 transcription<sup>225</sup>. ABCA2 facilitates the glycosylation and trafficking of Nicastrin. B-arrestin 1 is required for the assembly of the nicastrin-APH-1 subcomplex. G-protein-coupled receptor 3 and  $\beta$ -arrestin 2 confer

an effect on the cellular location of active gamma-secretase complexes (G-protein-coupled receptor 3 also appears able to increase overall complex formation)<sup>226; 228</sup>. GSAP and NDUFH7 appear to be required for effective A $\beta$  processing<sup>230; 231</sup>.

Opposed to these apparent enzyme activity enhancers a number of proteins have recently been reported to suppress complex activity. These comprise retention in endoplasmic reticulum 1 (Rer1), serum and glucocorticoid-induced protein kinase 1 (SGK1), retinoic acid, calreticulin, tubulin polymerisation promoting protein, tetraspanins, TMP21 and CD147<sup>222; 231-237</sup>. Retention in endoplasmic reticulum 1 (Rer1) is a negative regulator of gamma-secretase complex assembly. Rer1 competes with A $\Phi$ -1 for binding to the trans-membrane domain of nicastrin, thus preventing subcomplex formation and nicastrin release from the ER<sup>222</sup>. It can also prevent PEN-2 from leaving the ER<sup>221</sup>. Synoviolin appears to regulate the proteosomal and lysosomal degradation of Rer1 via ubiquitination. Over-expression of synoviolin increases complex formation, enzyme activity and A $\beta$  formation<sup>238</sup>. SGK1 phosphorylates nicastrin to induce its proteosomal or lysosomal degradation, thus reducing complex formation<sup>232</sup>. Over-expression of retinoic acid or calreticulin reduces A $\beta$  levels<sup>233; 234</sup>. Knockdown or silencing of tubulin polymerisation promoting protein, tetraspanins (CD9, CD81, EWI-F or CD98hc), TMP21 and CD147 increases A $\beta$  production<sup>231; 235-237</sup>.

Given the link with Alzheimer's disease the vast majority of research in this field has focussed on the interaction of gamma-secretase with APP. Many of the above interactors only appear to affect A $\beta$  processing (carboxypeptidase-like activity) rather than  $\epsilon$ -cleavage (endopeptidase activity) and as such may confer specificity towards APP as a substrate. Regardless, this work provides a unique insight in to the variable machinery that can associate with the complex to influence enzyme activity and



demonstrates that some of these proteins may differentially regulate the endopeptidase and carboxypeptidase-like activity of gamma-secretase.

Two large studies have recently been performed to define the gamma-secretase interactome. The first of these used *PSEN1* and *PSEN2* knock-out mouse embryonic fibroblasts<sup>235</sup>. By transfecting flag-tagged *PSEN1* and *PSEN2* they were able to pull down active complexes and identify associated proteins via SDS-PAGE and mass spectrometry. The second analysed TAP-tagged presenilin from mouse brains and HEK293 cells in a similar manner<sup>239</sup>. Both studies identified the three other integral complex components NCSTN, PEN-2 and APH-1 along with 134 other proteins. These comprised families of proteins involved in protein trafficking, cell adhesion, cell signalling, transporter proteins, proteins involved in glycosylation, proteins integral to effective protein folding and proteins integral to the tetraspanin web. Ten proteins were duplicated across the two studies (Table 1.1).

**Table 1.1. Proteins that have been shown to interact with PSEN1 in two independent studies**

Function	Protein	Gene
Trafficking and organisation	Tmp21	<i>TMED10</i>
	VCP	<i>VCP</i>
Adhesion and signalling	Catenin alpha-1	<i>CTNNA1</i>
	Catenin beta-1	<i>CTNNB1</i>
Transporters	Na(+)/ K(+) ATPase1	<i>ATP1A1</i>
	Proteolipid protein 2	<i>PLP2</i>
	V-ATPase subunit D1	<i>ATP6V0D1</i>
	V-ATPase subunit A	<i>ATP6V1A</i>
Tetraspanin	PGRL	<i>IGSF8</i>
Protein folding	HSC70	<i>HSPA8</i>

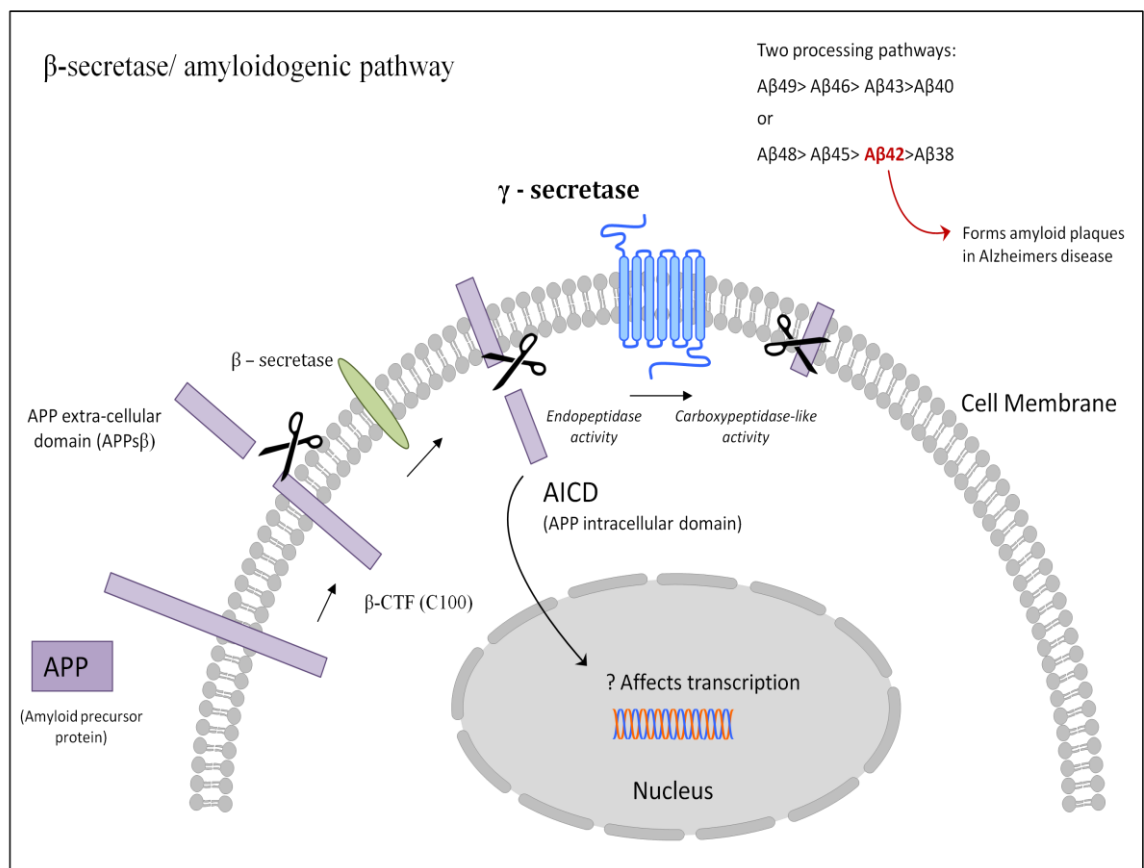
Whilst there was some disparity between the proteins identified in these studies they do provide an insight in to further complex interactors. Furthermore, the complex specificity of some of these proteins lends support to the growing evidence that subtle variation in complex composition (ie. PSEN1/ PSEN2, APH-1A/APH-1B) alters enzyme interactions and behaviour. For example, signal peptide peptidase (SPP) only appears to interact with *PSEN2* containing complexes whereas catenins, cadherins and plakophilins are more associated with *PSEN1*<sup>239</sup>.

### 1.3.5 Gamma-secretase processing of APP

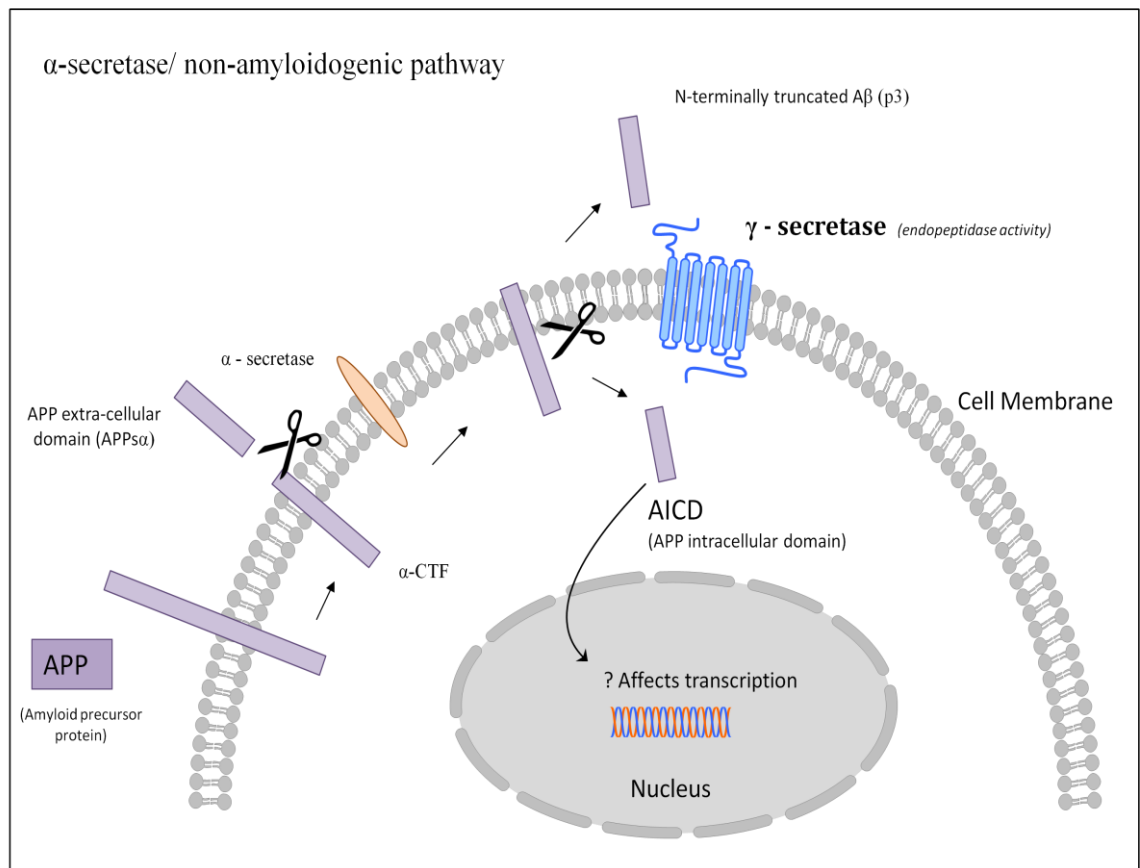
Gamma-secretase cleaves in excess of 30 type 1 membrane proteins however the mechanisms by which it processes APP and Notch have been studied in the most detail and provide an insight in to enzymatic function.

The biological role of APP is unclear however it appears to affect fibroblast growth patterns, synapse density and memory retention in animals and play a role in neurite outgrowth<sup>221; 222; 223; 224; 240</sup>. APP is produced in the ER, modified in the Golgi (glycosylated, phosphorylated and undergoes tyrosine sulphation) and a small proportion (10%) is transported to the cell membrane. The 695aa protein comprises a large ectodomain, a transmembrane region and an intracellular domain. The large ectodomain is initially cleaved by a  $\beta$ -secretase (BACE1)<sup>241</sup> or an  $\alpha$ -secretase (ADAM17, ADAM9, ADAM10, MDC-9, BACE-2)<sup>242</sup> which directs further processing down one of two routes, amyloidogenic or non-amyloidogenic pathways respectively (Figure 1.7, Figure 1.8).  $\beta$ -secretase cleavage leaves the transmembrane and intracellular domain of APP intact (100 peptide termed C100 protein). Gamma-secretase cleaves this protein (endopeptidase activity) resulting in the release of the APP intracellular domain (AICD). This initial cleavage leaves a portion of the protein within the cell membrane ( $\beta$ -amyloid) which is processed via a series of sequential cleavages (each removing 3 - 4aa's, carboxypeptidase-like activity). This results in the release of  $\beta$ -amyloid peptides of different lengths. Depending upon the length of the membrane stub left following AICD cleavage (can differ by 1aa, A $\beta$ 48 or A $\beta$ 49), two different paths of A $\beta$  processing can be followed generating peptides of different lengths (see Figure 1.7)<sup>178</sup>. Most secreted A $\beta$  is of the A $\beta$ 40 length (90%), smaller amounts of A $\beta$ 42 and A $\beta$ 38 are released. The longer A $\beta$ 42 peptide is more amyloidogenic and prone to forming plaques<sup>243</sup>.

$\alpha$ -secretase cleavage of APP leaves a shorter transmembrane and intracellular domain ( $\alpha$ -CTF) which is cleaved by gamma-secretase to form AICD and an N-terminally truncated A $\beta$  (p3)<sup>240</sup>. The N-terminally truncated A $\beta$  does not undergo any further processing by gamma-secretase. The biological role of AICD is unclear, however like NICD it may re-locate to the nucleus and affect gene transcription<sup>244</sup>. The biological role of A $\beta$  is unknown.



**Figure 1.7.  $\beta$ -secretase/  $\gamma$ -secretase processing of amyloid precursor protein (APP).** The APP ectodomain is initially cleaved by  $\beta$ -secretase, releasing the APP extra-cellular domain (APPs $\beta$ ) and leaving a portion within the cell membrane ( $\beta$ -CTF/ C100). C100 is then cleaved by  $\gamma$ -secretase to release the APP intracellular domain. This leaves an intramembranous stub which is sequentially cleaved (“processed”) to form shorter and shorter A $\beta$  fragments. The shorter fragments are eventually released from the membrane. These comprise A $\beta$ 38, a $\beta$ 40, a $\beta$ 42 and A $\beta$ 43. The longer peptides, particularly A $\beta$ 42, are more prone to aggregation and can form the amyloid plaques that are found in Alzheimer’s disease.

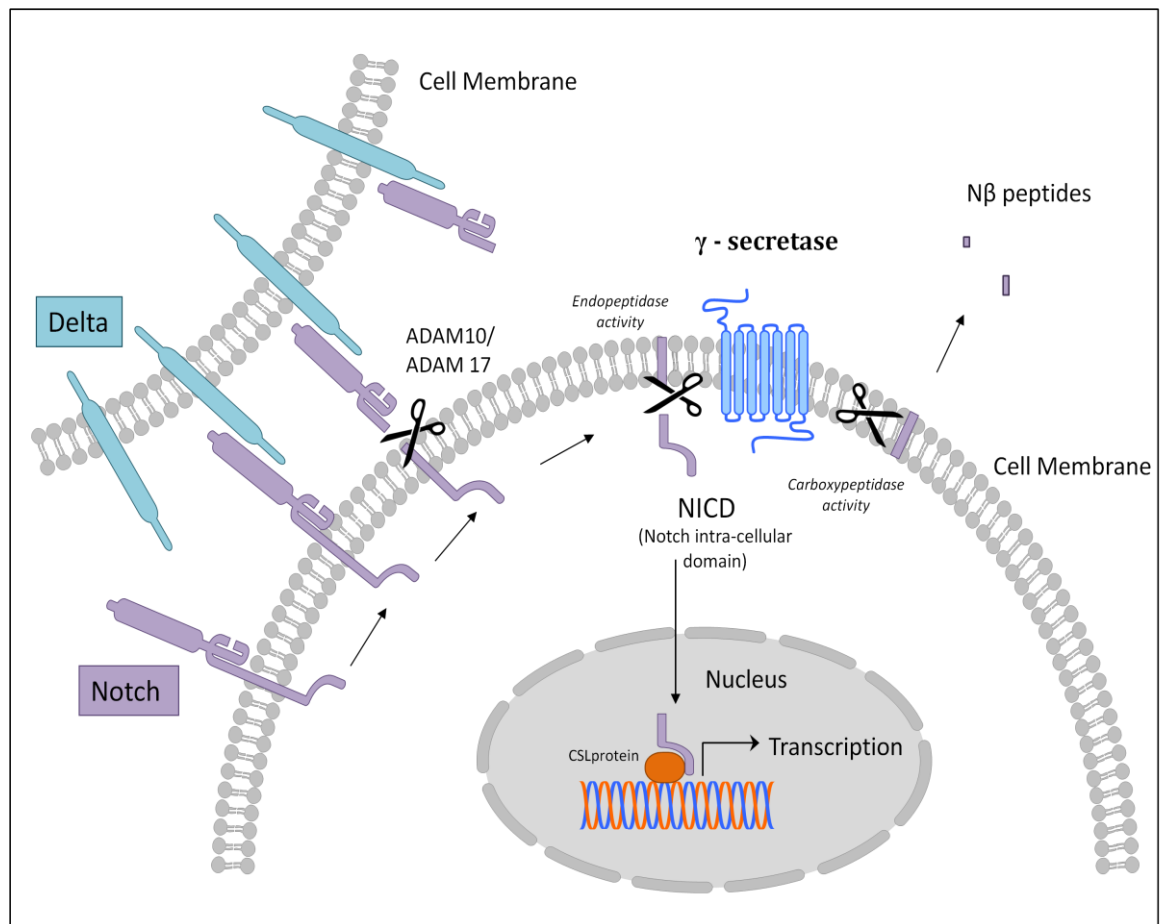


**Figure 1.8.  $\alpha$ -secretase/  $\gamma$ -secretase processing of amyloid precursor protein (APP).** The APP ectodomain is initially cleaved by  $\alpha$ -secretase, releasing the APP extra-cellular domain (APPs $\alpha$ ) and leaving a portion within the cell membrane ( $\alpha$ -CTF).  $\alpha$ -CTF is then cleaved by  $\gamma$ -secretase to release the APP intracellular domain (AICD) and the N-terminally truncated A $\beta$  protein. This mechanism of APP processing does not result in the formation and deposition of short A $\beta$  peptides, hence it is also referred to as the non-amyloidogenic pathway.

### 1.3.6 Gamma-secretase processing of Notch

Notch signalling controls cell fate decisions, differentiation, proliferation, apoptosis and cell and tissue patterning across the diverse array of tissue types required for normal human development<sup>245</sup>. Notch is synthesised in the ER as a 300kDa precursor protein and then transported to the Golgi where it is cleaved by furin-like convertase. The two cleavage products are non-covalently linked to form the notch heterodimer receptor which then undergoes sequential glycosylation prior to incorporation in to the cell membrane. EGF-repeats within the Notch ectodomain bind to the N-terminal ligand

domain of CSL ligands (Delta, Serrate and Jagged) expressed on the surface of other cells. The associated conformational change enables cleavage of the Notch ectodomain by ADAM10, ADAM 17, Kuz or SUP-17 leaving the Notch extracellular truncation (NEXT) in the membrane. The intracellular portion (Notch intracellular domain, NICD) is subsequently cleaved by gamma-secretase (endopeptidase activity), re-locates to the nucleus and affects gene transcription by binding with CSL, Mastermind (Mam) and co-activators<sup>245; 246</sup> (Figure 1.9). The remaining membrane stub is sequentially cleaved (as with A $\beta$  in the case of APP processing) to form N $\beta$  peptides of varying lengths<sup>247</sup>. The functional consequences of these peptides are unclear.



**Figure 1.9. Gamma-secretase processing of Notch receptors.** The Notch receptor sits in the cell membrane and binds to ligands on opposing cell membranes. Binding induces a conformational change in the receptor that allows ADAM10 or ADAM17 to cleave the extracellular portion of the receptor. This leaves the Notch extracellular truncation (NEXT) in the membrane which is in turn cleaved by gamma-secretase (endopeptidase activity) to release the notch intracellular domain (NICD). This re-locates to the nucleus and can affect gene transcription. The remaining intramembranous stub is sequentially cleaved (“processed”, carboxypeptidase-like activity) by gamma-secretase resulting in the release of Nβ peptides.

### 1.3.7 *PSEN1* and *PSEN2* mutations reported in familial Alzheimer’s disease (FAD)

The gamma-secretase complex became a focus of intense investigation following the discovery of heterozygous mutations in *PSEN1* and *PSEN2* in familial Alzheimer’s Disease (FAD)<sup>169; 248</sup>. Over 100 mutations have now been reported in presenilin and these are almost exclusively heterozygous missense variants

(<http://www.molgen.ua.ac.be/ADMutations>). Whilst some of these mutations affect the initial endopeptidase cleavage of APP they unanimously appear to affect the carboxypeptidase-like activity of gamma-secretase<sup>178</sup>. All mutations appear to increase the A $\beta$ 42: A $\beta$ 40 ratio<sup>178; 249</sup> by increasing A $\beta$ 42 production, reducing A $\beta$ 40 production or via a combination of the two mechanisms. This appears to be due to an impairment of the fourth and final cleavage step in each of the two A $\beta$  processing pathways (A $\beta$ 43 > A $\beta$ 40 and A $\beta$ 42 > A $\beta$ 38, see Figure 1.7) increasing the number of longer peptides produced. This may occur as a result of the peptides being prematurely released before the final cleavage can occur<sup>178</sup>. These longer peptides (A $\beta$ 42 and A $\beta$ 43) are more prone to aggregate (more amyloidogenic) and may be neurotoxic<sup>250-252</sup>. The functional consequences of these mutations therefore support the theory that the cerebral deposition of amyloid in the form of amyloid plaques drives the development of Alzheimer's disease (the amyloid hypothesis)<sup>243; 253</sup>. Another interesting observation is that mutations have been reported in APP in Alzheimer's disease which appear to preferentially direct processing down the A $\beta$ 42/ A $\beta$ 38 pathway rather than the A $\beta$ 43/ A $\beta$ 40 route. This effect was also observed with some *PSEN1* variants<sup>178</sup>. Subtle changes in the ratios of these A $\beta$  peptides significantly alter their biological behaviours, both in terms of aggregation and toxicity<sup>254</sup>. The ability of A $\beta$ 40 to stabilise the more amyloidogenic peptides and thus negate their aggregatory and toxic potential explains why a shift towards the A $\beta$ 42/A $\beta$ 38 pathway may be detrimental (A $\beta$ 38 does not confer the same level of stability as A $\beta$ 40)<sup>178</sup>. Rather than considering the total amount of A $\beta$  peptides present as being the most important factor in disease pathogenesis, the recent data suggests that it is the ratio of the individual peptides that plays a more crucial role<sup>255</sup>. Gamma-secretase inhibitors (eg. Semagacestat) have recently been trialled in Alzheimer's but conferred a negative impact on cognition and were associated with an array of what were thought to be Notch-related side effects including skin rashes.



Gamma-secretase modulators have more recently been developed (one being a non-steroidal anti-inflammatory drug (NSAID) derivative) which appear to selectively modulate the carboxypeptidase activity of gamma-secretase, increasing the fourth cleavage in each of the two processing pathways ( $A\beta_{42}/A\beta_{38}$  pathway >  $A\beta_{43}/A\beta_{40}$ )<sup>178</sup>.

## 1.4 Aims and objectives

The overall aim of this programme of research was to understand the genetic architecture of HS, the ultimate goal being to clarify the molecular mechanisms involved in disease pathogenesis. The principle objectives were three-fold:

- 1/ Identify genetic variants that underlie susceptibility to familial HS.
- 2/ Functionally characterise the identified mutations to elicit the molecular mechanisms by which they cause HS.
- 3/ Determine the contribution of any identified disease gene(s) to other familial and sporadic cases of HS

Shortly after this project started mutations were reported in the gamma-secretase genes in a small number of Chinese HS kindreds. Objective one therefore changed accordingly and became; assess the involvement of mutations in *NCSTN*, *PSENEN* and *PSEN1* in familial cases of HS.

The first part of this thesis (Chapter 3) explores the involvement of *NCSTN*, *PSEN1* and *PSENEN* mutations in seven British multiplex kindreds with HS. The second part (Chapter 4) describes the functional characterisation of the identified mutations. This section investigates the effect of each mutation on the relevant gene transcript, protein and on gamma-secretase enzyme activity. The histological distribution of these proteins in both healthy and diseased skin was assessed via immunohistochemistry. The third part (Chapter 5) explores the involvement of *NCSTN*, *PSENEN* and *PSEN1* mutations in a large cohort of affected individuals recruited from a tertiary referral hidradenitis clinic. The aim of that study was to determine the prevalence of mutations in these genes in the general disease population. Mutations in *NCSTN*, *PSENEN* and *PSEN1* only appeared to account for a minority of HS cases thus the fourth part of the thesis

(Chapter 6) describes the pursuit of further disease genes in unresolved familial cases and in well-defined phenotypic subgroups using a combination of gene mapping techniques including linkage analysis and next generation sequencing. The final part (Chapter 7) examines the clinical phenotype of all of the individuals found to harbour mutations in the gamma-secretase genes.

## **Chapter 2: Materials and Methods**

### **2.1 Patient and control cohorts**

#### **2.1.1 Regulatory and ethical approval**

Prior to proceeding with patient recruitment the following steps were completed:

1/ Mandatory training (Good Clinical Practice (GCP) course)

2/ Designing a study protocol

3/ Drafting study documentation

3/ Application for ethical approval

4/ Application for local Research and Development approval

I was actively involved in designing the study protocol, writing the study documentation and preparing the ethics application for this project. I attended all relevant research ethics committee meetings and made any necessary amendments to the study protocol and documents prior to commencement. This study was approved by East London Research Ethics Committee 2 (09/H0704/50) and London Stanmore Research Ethics Committee (11/LO/0966) and conducted in accordance with Declaration of Helsinki principles. All approved study documentation (protocol, letters of invitation, patient information leaflets, consent documents, clinical research forms and advertisements to recruit control subjects) can be seen in Appendix 2.

#### **2.1.2 Clinical definition of Hidradenitis Suppurativa**

Based on diagnostic criterion previously devised for the purposes of research and formal clinical diagnostic criterion agreed at a recent international HS conference, the diagnostic criteria applied in this study were as follows:

A history of five or more painful or discharging nodules, abscesses or sinus tracts in the axillae, infra-mammary folds, groin, perineal region, perianal region or buttocks.

### **2.1.3 Patient recruitment**

All probands were recruited from tertiary referral hidradenitis suppurativa clinics at either St. John's Institute of Dermatology, Guy's and St. Thomas' Hospitals, London or St. George's Hospital, London. All patients were presented with a letter of invitation and an information leaflet about the study. If they were happy to take part they were consented and a clinical history and examination was performed as detailed in the patient clinical research form, Appendix 2. If they were deemed to fulfil the diagnostic criteria detailed above they were asked to donate a blood (18ml in EDTA) or saliva sample (Oragene sample collection kit, DNA Genotek Inc., Kanata, Ontario, Canada) for DNA and RNA analysis. Selected individuals were also asked to donate two 6x6 mm skin biopsies to provide tissue for histological analysis and fibroblasts for culture. The two skin biopsies were harvested from an affected and unaffected area of the axilla.

### **2.1.4 Recruitment of family members**

Those individuals who presented with a family history of HS were asked to discuss the project with their respective family members. Any relatives who were interested in the study were sent a letter of invitation and an information leaflet. If, having read those documents, they were happy to be involved in the study then they were asked to return a reply slip and a formal appointment was arranged at St. John's Institute of Dermatology. Individuals were consented, questioned, examined, had their affection status determined and asked to provide samples as documented above. For the minority of individuals who could not attend the outpatient department arrangements were made for a telephone interview and postal collection of a saliva sample, or a patient visit (Dermatologist accompanied by a Dermatology Specialist Nurse).

A total of 300 individuals were consented and recruited to the study. These comprised 237 probands and 63 family members from 10 families (40 with, 23 without HS).

#### **2.1.5 Recruitment of healthy volunteer subjects (controls)**

Healthy volunteers were recruited following the release of a University-wide advertisement (Appendix 2). All volunteers were seen in the St. John's Institute of Dermatology outpatient department where they were given a study invitation and information leaflet. If they were happy to take part they were subsequently consented to the study and provided a clinical history and underwent an examination as documented in the volunteer CRF (Appendix 2). Inclusion criteria were as follows:

No history of comedones, papules, pustules, nodules, abscesses, sinus tracts or hypertrophic scars in the post-auricular region, axillae, infra or intra-mammary folds, abdominal fold, supra-pubic region, groin, perineal region, perianal region, buttocks or medial thighs. No history of an associated disorder of follicular occlusion or acne vulgaris and no family history of HS.

All healthy volunteers who fulfilled the above criteria were asked to provide a blood sample for DNA or RNA analysis and 2 skin biopsies (performed exactly as documented above for the patients). Volunteers were age and sex-matched as closely as possible to the patients against whom they were being compared and were paid for their time. A total of 6 healthy volunteers were recruited to the study.

#### **2.1.6 Recruitment of patients with nodulocystic acne**

Patients with acne vulgaris have recently been recruited from general dermatology clinics at St. John's Institute of Dermatology for a genome wide association study of acne vulgaris (study conducted in accordance with Declaration of Helsinki principles and approved by the St. Thomas' Hospital Research Ethics Committee (05/Q0702/114)). Study criteria stated that all individuals had to have moderate-severe

acne, as evidenced by a Leeds Grading Score of >5 or a clinical report of moderate disease and previous treatment with isotretinoin (examined by a Dermatologist). In each case, the clinical sub-type of acne (eg. nodulocystic acne) was documented on the clinical research form. All patients were asked to donate a blood or saliva sample for DNA analysis. Forty eight patients with nodulocystic acne vulgaris were identified from this patient cohort and their DNA was used in this study.

## 2.2 Antibodies used

The following table (Table 2.1) lists the antibodies used in this study:

**Table 2.1. Antibodies used**

<b>Primary Antibody</b>	<b>Supplier</b>	<b>Dilution</b>	<b>Order number</b>
<b>Rabbit polyclonal nicastrin</b>	Abcam (Cambridge, UK)	1:400 WB  1:200 IHC	Ab24741
<b>Rabbit polyclonal nicastrin</b>	Abcam	1:750	Ab122969
<b>Rabbit polyclonal PEN-2</b>	Abcam	1:200 WB  1:20 IHC	Ab18189
<b>Mouse monoclonal anti-Flag</b>	Sigma (Poole, UK)	1:1000	F3165
<b>Rabbit polyclonal anti-<math>\beta</math>-actin</b>	Cell Signalling Technology (Hitchin, UK)	1:1000	4967S
<b>Rabbit polyclonal anti-myeloperoxidase</b>	DAKO UK Ltd (Ely, UK)	1:5000	A0398
<b>Rabbit polyclonal anti-CD3</b>	DAKO UK Ltd	1:500	A0452
<b>Mouse monoclonal anti-L26</b>	DAKO UK Ltd	1:100	M0755
<b>Mouse monoclonal anti-CD68</b>	Abcam	1:1	Ab844



<b>Secondary Antibody</b>	<b>Supplier</b>	<b>Dilution</b>	<b>Order number</b>
<b>Donkey anti-rabbit IgG horseradish peroxidise linked-antibody</b>	GE Healthcare UK Ltd. (Amersham, UK)	1:10000	NA934V
<b>Goat polyclonal anti-mouse IgG horseradish peroxidase-linked antibody</b>	DAKO UK Ltd	1:10000	P0447
<b>Biotinylated Goat anti- rabbit IgG</b>	Vector Laboratories (Peterborough, UK)	1:200	BA-1000

WB = Western blotting, IHC = Immunohistochemistry

## 2.3 Buffers used

The following table (Table 2.2) lists the common buffers used in this study.

**Table 2.2. In-house buffers**

<b>Buffer</b>	<b>Recipe</b>
<b>6M guanidium hydrochloride</b>	573.18g guanidium hydrochloride (Sigma) in 1l dH <sub>2</sub> O
<b>7.5M ammonium acetate</b>	115.62g ammonium acetate (Sigma) in 200ml dH <sub>2</sub> O
<b>20% sodium sarkyl</b>	100g N lauryl sarcosine (Sigma) in 500ml dH <sub>2</sub> O
<b>96% ethanol</b>	48ml pure ethanol (VWR International Limited, Lutterworth), 2ml dH <sub>2</sub> O
<b>95% ethanol</b>	47.5ml pure ethanol (VWR International Limited), 2.5ml dH <sub>2</sub> O
<b>70% ethanol</b>	35ml pure ethanol (VWR International Limited), 15ml dH <sub>2</sub> O
<b>Tris-EDTA (TE) buffer</b>	10ml 100x TE buffer (Sigma) diluted in 1l dH <sub>2</sub> O
<b>3M sodium acetate (pH 5.5)</b>	81.64g sodium acetate (Sigma) in 200ml dH <sub>2</sub> O
<b>10xTBE</b>	890 mM Tris-base, 890 mM Boric acid and 20mM EDTA (pH 8.0)
<b>1xTBE</b>	100ml 10 x TBE, 900ml dH <sub>2</sub> O
<b>10xTBS</b>	30g Tris, 44g NaCl, 350ml dH <sub>2</sub> O, pH adjusted to 7.6 and

	volume topped up to 500ml with dH <sub>2</sub> O
<b>1xTBS</b>	50ml 10x TBS, 450ml dH <sub>2</sub> O
<b>TBS-T</b>	100ml 10x TBS, 900ml dH <sub>2</sub> O, 1ml Tween (Sigma)

## **2.4 Molecular biology techniques**

### **2.4.1 DNA extraction from blood (leucocytes)**

Blood samples were placed in 50ml centrifuge tubes and mixed with 30ml ice cold dH<sub>2</sub>O to induce erythrocyte cell lysis. Samples were spun at 1250g for 20 minutes at 4°C and the supernatant was discarded. The leukocyte-rich cell pellet was solubilised in 30ml ice cold 1% IGEPAL (SIGMA) and the sample was vortexed and spun at 1250g for 20 minutes at 4°C (Rotanta 460R, Hettich, Salford, UK). The supernatant was discarded and DNAase inhibition and denaturation of the leukocyte lysates was achieved by adding 7ml of filtered 6M guanidium hydrochloride (SIGMA), 0.5ml 7.5M ammonium acetate (SIGMA), 0.5ml 20% sodium sarkosyl (SIGMA) and 150 µl proteinase K (SIGMA) (10mg/ml). Samples were vortexed and incubated at 60°C for 90 minutes. DNA was precipitated by adding 17ml 96% ethanol (EtOH) and gently mixing the sample. DNA was spooled out on to the tip of a Pasteur pipette. Pipette tips were placed in 1 ml of tris-EDTA (TE) buffer (SIGMA) in 10 ml centrifuge tubes. The DNA was detached from the pipette tips and re-dissolved overnight at 5°C. DNA was re-precipitated with 100µl 3M sodium acetate (pH 5.5, SIGMA) and 2.5ml ice cold 96% EtOH. DNA was spooled out as before, allowed to dry and re-dissolved in 500µl of TE buffer at 5°C. The DNA concentration was measured on the Nanodrop Spectrophotometer (ND-1000, Thermo Scientific, Loughborough, UK) and samples were stored at -80 °C.

### **2.4.2 DNA extraction from saliva**

Saliva was collected as per manufacturer protocol in Oragene saliva vials (Oragene sample collection kit, DNA Genotek Inc., Ontario, Canada). Vials were incubated at 50°C for 60 minutes and each sample was divided in to 4 1ml aliquots. 40µl PrepIT L2P (commercial proprietary reagent, contents not disclosed by company, DNA Genotek Inc.) was added to each tube and the tubes were inverted and incubated on ice for 10

minutes. Samples were spun at 15000g (Centrifuge model 5415R, Eppendorf UK Limited, Stevenage, UK) for 3 minutes at room temperature and the supernatant from all four tubes was combined in one 15ml centrifuge tube. An equal volume of 95% EtOH was added to precipitate DNA and samples were inverted and allowed to stand for 10 minutes at room temperature. Each sample was spun at 1100g for 10 minutes at room temperature and the supernatant was discarded. A further spin at 1100g for 2 minutes was performed and all excess supernatant was removed. The pellet was allowed to dry and was re-dissolved in 500µl TE buffer overnight at 5°C. The DNA concentration was measured on the Nanodrop (Thermo Scientific) and samples were stored at -80 °C

### 2.4.3 Polymerase chain reaction (PCR)

Primers were designed to amplify all coding exons and associated splice sites of *NCSTN* (17 exons), *PSENI* (10 exons), *PSENE1* (3 exons) (primer sequences listed in Appendix 3). All primers were ordered from Eurofins MWG Operon, Ebersberg, Germany. The PCR reaction mix totalled 10µl (components shown in Table 2.3):

**Table 2.3. PCR reaction mix**

<b>DNA</b>	1 µl
<b>10 µM forward primer</b>	0.3 µl
<b>10 µM reverse primer</b>	0.3 µl
<b>2 µM dNTPs (ABgene Limited., Epsom, UK)</b>	1 µl
<b>10x polymerase buffer (ABgene)</b>	1 µl
<b>5 U/ µl Taq polymerase (ABgene)</b>	0.1 µl
<b>dH<sub>2</sub>O</b>	6.3 µl

Every PCR comprised a denaturation step (96°C for 5 minutes), 35 replication cycles (96°C for 30 seconds, 56 - 62°C (depending on primer requirements) for 30 seconds

(annealing) and 72°C for 30 seconds (elongation)) and a final clean-up step (72°C for 10 minutes). All PCR cycles were performed on the G-Storm GS4 Thermal cycler (Labtech International Ltd., Ringmer, UK).

#### **2.4.4 Analysis of PCR products by agarose gel electrophoresis**

1g of agarose (Sigma) was added to 100ml of 1 x TBE buffer and heated in the microwave for 2 minutes until fully dissolved. 5 µl ethidium bromide (Sigma) was then added to the cooled agarose solution and it was poured out and allowed to set. DNA samples were mixed with 5 x loading buffer (Bioline, London, UK) and run on the gel at 100mA for 20-30 minutes. The amplified products were analysed on a UV transilluminator (GelDoc-It Imaging System, UVP, Cambridge, UK).

#### **2.4.5 Direct nucleotide sequencing**

All PCR products were purified using ExoSAP (Affymetrix, High Wycombe, UK). 2 µl DNA was added to 0.26 µl of ExoSAP and 5.74 µl dH<sub>2</sub>O and incubated at 37°C for 30 minutes and 80°C for 15 minutes. The product was sequenced using a Big-dye terminator kit (Life Technologies Limited, Paisley, UK). 3.5 µl of product was added to 1.25 µl Big-dye sequence buffer (Life Technologies), 0.25 µl Big-dye 3.1 (Life Technologies) and 0.25 µl forward primer. A further 3.5 µl of product was added to 1.25 µl Big-dye sequence buffer, 0.25 µl Big-dye 3.1 and 0.25 µl reverse primer. Both mixes were thermo-cycled at 96°C for 30 seconds, 50°C for 15 seconds and 60°C for 60 seconds for a total of 30 cycles. Samples were further purified via ethanol precipitation. 26 µl of precipitation solution (50ml EtOH, 2ml 3M NaOAc pH4.6 (SIGMA)) was added to each sample. Samples were allowed to stand for 10 minutes then spun at 3000rpm for 30 minutes at room temperature. 100 µl 70% EtOH was added to each sample and samples were spun at 3000rpm for 10 minutes at room temperature and then left to dry for 10 minutes. 10 µl HiDi Formamide (Life Technologies) was added to each reaction and the products were heat denatured (incubated at 90°C for 2 minutes).

All sequencing reactions were run in 96 well plates on the ABI 3730XL Automated Sequencer (reagents supplied by Applied Biosystems, Life Technologies). Results were analysed using Sequencher software (Gene Codes Corporation, Michigan, USA). All sequence variants were detected via visual inspection.

#### **2.4.6 Genotyping using Exome chip**

Genotyping was performed using the Illumina Infinium HD Assay Ultra kit and Infinium HumanExome 12v1.1 chip following the published protocol (Illumina, Little Chesterford, UK). All samples were prepared using automation (Freedom Evo, Tecan Group Ltd., Reading, UK). 200ng DNA was denatured and neutralised. The denatured DNA was isothermally amplified overnight (increasing total amount of DNA by several thousand-fold) and fragmented via a controlled enzymatic process using end-point fragmentation. DNA was precipitated by adding isopropanol and pelleted with centrifugation. The precipitated DNA was re-suspended in hybridisation buffer and loaded on to a BeadChip, divided by an IntelliHyb seal. The chip was incubated overnight in a hybridisation oven resulting in the hybridisation of DNA fragments to locus specific 50-mers (all covalently linked to one of up to 300000 bead types). The Beadchip was washed to remove any unhybridised or non-specifically hybridised DNA. The bound DNA was used as a template to add a single labelled base to the oligos on the BeadChip (single-base extension). The oligos are designed in such a way that the base following the sequence of bases forming the oligo is the one of interest (single base polymorphism). Thus the identity of the single base extension (A, T, C or G) represents the genotype call for that particular SNP in that sample. These single bases are labelled so when the Beadchip is placed in a BeadArray Reader the identity of these bases can be identified (a laser excites the fluorophore of the single-base extension product and the scanner records high resolution images of the light emitted from the fluorophores).

#### **2.4.7 Exome chip data analysis to define regions of linkage within pedigrees**

Data for around 5000 single nucleotide polymorphisms (SNPs), representing a linkage panel ([www.illumina.com/exome](http://www.illumina.com/exome)), was collated for individuals within any one multiplex kindred and analysed using MERLIN (Multipoint Engine for Rapid Likelihood Inference) software<sup>256</sup>. An affected only analysis was performed given the variable age of onset observed in HS and an autosomal dominant model was adopted as the disease segregated as an autosomal dominant trait in the pedigrees studied.

#### **2.4.8 Whole exome sequencing method**

Whole exome capture was performed via in-solution hybridisation followed by massively parallel sequencing. In-solution hybridisation was performed using the SureSelect All Exon 50Mb Target Enrichment System (Agilent, Wokingham, UK). The first 16 samples were manually prepared and the remaining 52 samples were prepared using automation (Agilent Bravo, Agilent BenchCel 4R, Agilent PlateLoc, Agilent). 3µg of genomic DNA was sheared to create DNA fragments with a maximum length of 150 – 200bp. These fragments were subjected to end-repair, blunting the ends of each fragment, and 5'-phosphorylation. Adenine bases were added to the 3' end to create an overhang. Adaptors were ligated to both ends of each fragment and unligated adaptors were removed. The adaptor-ligated library was then amplified (4 PCR cycles) to form the prepped library. Biotinylated RNA library "baits" (Agilent Whole Exome XT Baits, Agilent), designed to bind to coding regions, were hybridized to the prepped genomic sample (24 hour hybridization) and bound DNA fragments were harvested using Streptavidin coated magnetic beads which selectively bind to the biotinylated RNA "baits" (bead capture or hybrid capture selection). Unbound DNA fragments were discarded. DNA fragments were eluted from the Streptavidin coated beads and the biotinylated RNA "baits" bound to the DNA fragments were digested. The DNA was



amplified via PCR (12 cycles) and index tags were added. Samples were pooled and sequenced on the HiSeq 2000 (Illumina) with 100bp paired end reads.

#### **2.4.9 Whole exome sequencing data preparation**

Sequence reads were aligned to the reference genome (hg19) with Novoalign (Novocraft Technologies, Kuala Lumpur, Malaysia). Reads mapping to multiple locations and duplicate reads (resulting from PCR clonality or optical duplicates) were excluded from downstream analysis. The depth and breadth of sequence coverage was calculated with custom scripts and the BedTools package<sup>257</sup>. More than 4.1Gb of sequence was generated for each individual and an average of 89.3% of the coding bases of the RefSeq-defined exome were represented by at least 20 reads. Single nucleotide substitutions and small insertion/ deletion variants were identified and quality filtered using the SamTools software package<sup>258</sup> and in-house software tools. Annovar<sup>259</sup> was used to annotate variants according to genes and transcripts. All variants were filtered against dbSNP137, NHLBI Exome Sequencing Project Exome Variant Server (EVS, <http://evs.gs.washington.edu/EVS/>), the 1000 Genomes Project database<sup>260</sup> and over 600 high quality in-house control exomes (sequenced and analysed by the method described above) to determine their novelty. Copy number variation was identified and quality filtered using the Exome Depth package<sup>261</sup> and annotated using BedTools<sup>257</sup>. All variants were filtered against the database of genomic variants (<http://dgvbeta.tcag.ca/dgv/app/home>) to determine novelty.

#### **2.4.10 RNA isolation from whole blood**

500µl of anti-coagulated (EDTA) whole blood was added to 1.3ml RNA later (Life Technologies) and mixed. RNA was extracted using a Ribopure Extraction Kit (Life Technologies, all reagents, spin columns and tubes were provided as part of this kit unless otherwise stated). The sample was centrifuged at 16000g for 1 minute at room temperature and the supernatant was removed. 800 µl lysis solution and 50µl sodium

acetate solution were added to the cell pellet to induce lysis and the sample was vortexed. 500µl of acid-phenol: chloroform was added and the sample was vortexed and left at room temperature for 5 minutes. The sample was centrifuged at 16000g for 1 minute at room temperature and the aqueous (upper) phase was transferred to a new 2ml tube. The sample was loaded to a gDNA eliminator mini spin column to remove DNA and spun at 8000g for 30s at room temperature. The column was discarded and 600µl of 100% EtOH was added to the remaining sample to precipitate RNA. 700µl of the sample was added to a filter cartridge assembly and spun at 16000g for 10 seconds. This was repeated with a further 700µl of sample and the flow-through was discarded. 700µl wash solution 1 was applied to the cartridge and spun at 16000g for 10 seconds, the through-flow was discarded. 700µl wash solution 2/3 was applied to the cartridge and spun at 16000g for 10 seconds and the through-flow was discarded. This step was repeated to thoroughly wash the RNA. The cartridge assembly was centrifuged at 16000g for 1 minute and transferred to a collection tube. 25 µl of elution buffer (pre-heated to 75 °C) was added and left for 20 seconds. The tube was spun at 16000g for 30 seconds to recover the RNA. RNA concentration was measured on the Nanodrop (Thermo Scientific) and the sample was stored at -80 °C.

#### **2.4.11 RNA isolation from cells**

Cell media was aspirated and the cells were washed with phosphate buffered saline (PBS, Gibco, Life Technologies). Cell detachment was induced with TrypLE Express 1x (Gibco, Life Technologies). Once the cells had detached they were mixed with an equal volume of medium containing serum (DMEM (1x) + GlutaMAX + 10% fetal bovine serum (FBS)) to inactivate the trypsin. The cells were transferred to a 15ml centrifuge tube and centrifuged at 1000rpm for 5 minutes at room temperature. The supernatant was aspirated and the cell pellet was suspended in 100-150ul RNA lysis solution (Life Technologies), transferred to a 1.5ml micro-centrifuge tube and frozen at

-80 °C. RNA was extracted from this sample using the Qiagen RNA Easy Plus Mini Kit (Qiagen, Crawley, UK) (all reagents, spin columns and tubes stated in the below method are provided with this kit unless otherwise stated). The sample was thawed on ice and the eppendorf was flicked twice. 600 µl of Buffer RLT Plus was mixed with 6 µl of β-Mercaptoethanol. 600µl of this mixture was added to the cell sample and the sample was vortexed for 3s to mix (to induce cell lysis). The lysate was homogenised by pipetting it in to a QIAshredder spin column placed in a 2ml collection tube and centrifuging at 15000rpm for 2 minutes. The spin column was discarded and the homogenised lysate was transferred to a gDNA Eliminator spin column and centrifuged at 8000rpm for 30s at room temperature to eliminate DNA from the sample. The spin column was discarded and 600 µl of 70% EtOH was added to the flow through to precipitate RNA and the sample was mixed well by pipetting. 700 µl of the sample was placed in an RNeasy spin column placed in a 2ml collection tube and centrifuged at 10000rpm for 15s at room temperature. The flow through was discarded. This step was repeated until of the sample had been run through the RNeasy spin column. 700 µl of Buffer RW1 was then added to the RNeasy spin column and centrifuged at 10000rpm for 15s at room temperature to wash the spin column membrane. The flow through was discarded. 500 µl of Buffer RPE was added to the RNeasy spin column and centrifuged at 10000rpm for 15s at room temperature to further wash the spin column membrane. The flow through was discarded. One final wash was performed by adding another 500 µl Buffer RPE to the RNeasy spin column and the column was centrifuged at 10000rpm for 2 minutes. The RNeasy spin column was placed in a new 2ml collection tube and centrifuged at 15000rpm for 1 minute to eliminate any remaining RPE buffer. The RNeasy spin column was then placed in a new 1.5ml collection tube. 30 µl RNase-free water was added to the spin column membrane and centrifuged at 10000rpm for 1

minute to elute the RNA. The RNA concentration was measured on the Nanodrop (Thermo Scientific) and the sample was stored at -80 °C.

#### **2.4.12 RNA interference**

35000 primary human fibroblasts were seeded in a 24-well tissue culture plate (Costar, Corning, New York, USA) well and grown until 30-50% confluent (roughly 24 hours, see maintenance of primary human fibroblasts, section 2.6.3). Cell media was removed and cells were washed with 1ml PBS (Gibco, Life Technologies) and the PBS was aspirated. A 20 $\mu$ M stock solution of *NCSTN* SiRNA (ON-TARGETplus SMARTpool siRNA J-008043-08, J-008043-07, J-008043-06, J-008043-05, Thermo Fisher Scientific, Loughborough, UK) was made up and diluted to 25nM in OptiMEM I Reduced Serum Medium (Gibco, Life Technologies). 1 $\mu$ l of Lipofectamine 2000 Transfection reagent (Invitrogen) was diluted in 50 $\mu$ l OptiMEM I Reduced Serum Medium and incubated at room temperature for 5 minutes. 50  $\mu$ l of this solution was combined with 50 $\mu$ l of the 25nM SiRNA solution, mixed and incubated at room temperature for 20 minutes. This mixture was added to the well and the well was topped up with an additional 400 $\mu$ l OptiMEM I Reduced Serum Medium. After 3 hours a further 500 $\mu$ l Cascade 106 media (Gibco, Life Technologies) + LSGS supplement kit (S-003-K, Gibco, Life Technologies) was added to the well and the cells were left incubating at 37°C for 48 hours. All SiRNA experiments were performed in triplicate. RNA was extracted from one well and protein was extracted from the remaining two wells. A further three wells of cells were seeded for each patient and not exposed to SiRNA but cultured in identical media without SiRNA so as to act as controls. RNA and protein were extracted from one well and two wells respectively as with the SiRNA-treated cells. RNA facilitated real time PCR (section 2.4.16) to determine if the knockdown had been successful and protein was used for immunoblotting as detailed in section 2.7.4.

#### 2.4.13 Verification of the specificity of an NCSTN antibody by RNA interference

RNA interference was performed to confirm the identity of the NCSTN protein band on immunoblotting. When RT-PCR had been performed and clarified that effective knockdown had occurred, protein from cells exposed to SiRNA and not exposed to SiRNA was subjected to electrophoresis and immunoblotted (see section 2.7.4). If the protein band observed does indeed represent NCSTN, then the band observed using protein harvested from knocked-down cells should be smaller and less dense than that observed from untreated cells.

#### 2.4.14 Reverse transcription polymerase chain reaction (RT-PCR)

RT-PCR was performed using the high capacity cDNA reverse transcription kit (Life Technologies). The PCR reaction components used are shown in

Table 2.4, each reaction totalled 20µl.

**Table 2.4. cDNA reverse transcription mix.**

<b>RNA</b>	10 µl
<b>10 x RT buffer</b>	2 µl
<b>25 x dNTP Mix (100mM)</b>	0.8 µl
<b>10 x RT random primers</b>	2 µl
<b>Multiscribe reverse transcriptase</b>	1 µl
<b>Nuclease-free H<sub>2</sub>O</b>	4.2 µl

The mix was incubated at 25°C for 10 minutes, 37°C for 120 minutes and 85°C for 5 minutes on the G-Storm GS4 Thermal cycler (Labtech International Ltd.). cDNA was stored at -20°C.

#### 2.4.15 cDNA PCR and direct sequencing

Exonic primers were designed to amplify a segment of *PSENEN* cDNA and the full length *NCSTN* cDNA (Table 2.5, ordered from Eurofins MWG Operon). PCR was performed as described in section 2.4.3 for *PSENEN* cDNA (but cDNA was used instead of DNA). Due to the size of the full length *NCSTN* transcript (2944kb) the PCR reaction mix was altered (see Table 2.6) and the reagents from the KOD Hot Start DNA Polymerase Kit (Novagen) were used. The PCR cycle was also different comprising a denaturation step (95°C for 5 minutes), 40 replication cycles (95°C for 30 seconds, 58°C for 30 seconds (annealing) and 70°C for 150 seconds (elongation)) and a final clean-up step (72°C for 10 minutes).

**Table 2.5. cDNA primers used for sequencing *NCSTN* and *PSENEN* cDNA**

Gene transcript	Forward primer	Reverse primer
NCSTN (spanning exons 1 -17)	CTACTAGCAGGTTTGTGCAG	AGAAGATGAGGATGCCGAAG
PSENEN (spanning exons 1-3)	ATGAACCTGGAGCGAGTGTC	ATGGTGATCCAGGAGGTGAG

**Table 2.6. PCR reaction mix used to amplify the full length *NCSTN* transcript**

<b>cDNA</b>	10 µl
<b>10 µM forward primer</b>	1.5 µl
<b>10 µM reverse primer</b>	1.5 µl
<b>25mM MgSO<sub>4</sub></b>	3 µl
<b>2 mM dNTPs</b>	5 µl
<b>10 x Buffer KOD Hot Start DNA Polymerase</b>	5 µl
<b>KOD Hot Start DNA polymerase (1U/ µl)</b>	1 µl
<b>PCR grade H<sub>2</sub>O</b>	23 µl

Direct sequencing was performed as described in section 2.4.5. Due to the length of the full length *NCSTN* transcript five pairs of overlapping primers (not including those used for the PCR step) were designed for sequencing (shown in Table 2.7) (Eurofins).

**Table 2.7. Primers used to sequence the full length *NCSTN* transcript**

<b>Gene transcript</b>	<b>Forward primer</b>	<b>Reverse primer</b>
NCSTN	TCCTTCGCCTTCTGTCTTTC	CTAAAGTCTTCATAAGCCAA
NCSTN	TATGGGCCAGAGTTTGCTCA	CGAGCTGCCAATGTAGTCAA
NCSTN	GGGACATTAAAGCCTGACGA	TGGTTCCGTACAGACTCATTTTT
NCSTN	GCAGCTCGAGGATGGTCTAC	TTAGGTCCTGCCTGAGGATA
NCSTN	GAACCAACTTCAGCGACACA	AGAAGATGAGGATGCCGAAG

#### **2.4.16 Gene expression assays**

Total *NCSTN* and *PSENEN* mRNA levels were assessed by multiplex real-time PCR using validated Taqman Gene Expression Assays (Life Technologies) as per the manufacturer's instructions. The PCR reaction mix comprised the components shown in

Table 2.8, totalling 20µl.

**Table 2.8. cDNA PCR reaction mix**

<b>cDNA</b>	1 µl
<b>20 x TaqMan Gene expression Assay (probe)</b>	1 µl
<b>2 x TaqMan Gene Expression Master Mix</b>	10 µl
<b>RNAase free H<sub>2</sub>O (Qiagen, Crawley, UK)</b>	7 µl
<b>Endogenous control</b>	1 µl

All PCR reactions were carried out in 96-well optical plates (ABgene) with optical adhesive lids. Samples were amplified in an Applied Biosystems 7900HT Sequence Detection System and all reactions were performed in triplicate. *NCSTN* and *PSENEN* amplification (Taqman probe details shown in Table 2.9) was performed under the following cycling conditions: 95°C for 10 minutes followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min. Target gene expression was normalized to that of *GAPDH* and/or *18sRNA* endogenous genes (Assay ID 4326317E, 4310893E respectively, Life Technologies). Results were analysed using SDS 2.4 (Life Technologies). Real-time data were expressed as mRNA expression fold changes using the comparative C<sub>T</sub> (2<sup>-ΔΔC<sub>t</sub></sup>) method<sup>262</sup>. The ΔΔC<sub>t</sub> value was calculated using the formula: ΔΔC<sub>t</sub> = (C<sub>T,Target</sub> - C<sub>Endogenous</sub>)<sub>Treated</sub> - (C<sub>T,Target</sub> - C<sub>Endogenous</sub>)<sub>Untreated</sub>.

**Table 2.9. Taqman probes used to study NCSTN and PSENEN transcript expression**

<b>Assay ID*</b>	<b>Gene</b>	<b>Target exon</b>	<b>Target sequence</b>
Hs00299716_m1	<i>NCSTN</i>	Exon 2	TGGCTGCCAGTCTTCAATTAGTGGA
Hs01033959_g1	<i>PSENEN</i>	Exon 1	AGTGAGCTCTCCTGGGGCGTGTTG

\* All assays were purchased from Applied Biosystems



#### **2.4.17 Multiplex Ligation-dependent Probe Amplification (MLPA)**

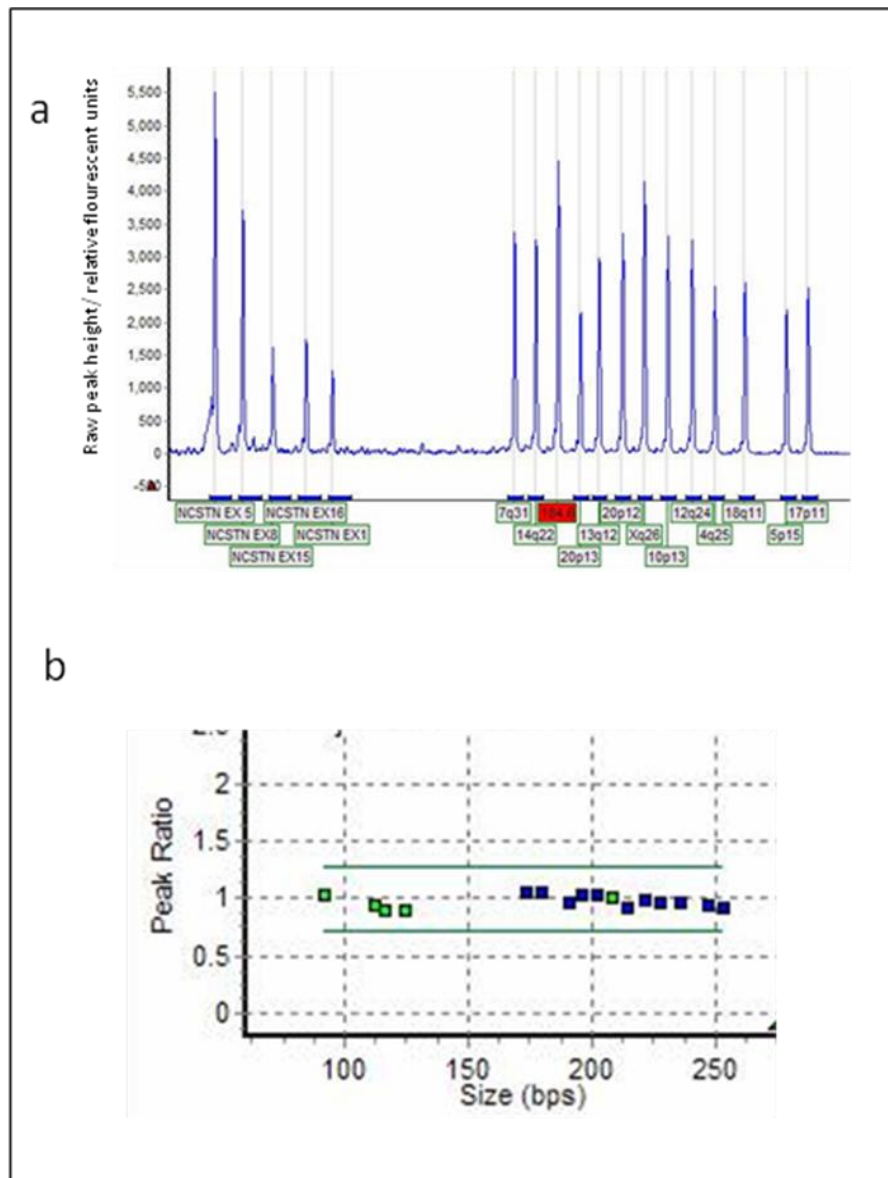
Assays were performed as per manufacturer protocol with SALSA MLPA kit P254 *PSENI* and SALSA MLPA kit P200-A1 (MRC-Holland, Amsterdam, Netherlands) which contains eleven reference probes and additional space for the insertion of user-designed custom probes. Custom probes were designed for all exons of *PSENI* (added to P254 *PSENI*) and *NCSTN* (added to P200-A1) as per MRC-Holland instructions (<http://mlpa.com>) and manufactured by Eurofins MWG Operon (probe sequences shown in Appendix 4). They were prepared and diluted according to the synthetic MLPA protocol (<http://mlpa.com>), as described in Hills *et al.* 2010<sup>263</sup>.

At least three control samples (unaffected individuals) and a no DNA control sample were run with every reaction. For analysis, 3ul of PCR product was added to 0.3ul Genescan 500 LIZ size standard (Applied Biosystems) and 15ul HiDi formamide prior to size separation on a 3730XL genetic analyser (Life Technologies) as per manufacturers instructions.

Data was analysed using GeneMarker v1.8 software (SoftGenetics, State College, USA). The maximum raw data peak height was set at 15000 relative fluorescent units (rfu) with a lower limit of 250 rfu (samples outside of these limits were failed). Peak height data from MRC-Holland P200 kits with added custom probes were analysed according to recommendations on data normalisation provided on the MRC-Holland website (<http://mlpa.com>), as described in Hills *et al.* 2010<sup>263</sup>. Peak height data from MRC-Holland P200 kits with added custom probes were analysed according to recommendations on data normalisation provided on the MRC-Holland website, as described in Hills *et al.* 2010<sup>263</sup>. Panels were designed to detect one or more custom probes per kit.

Copy number changes were displayed via application of the MLPA Ratio function (peak height ratios defined deletion as  $<0.71$ , duplication  $> 1.23$ ). Patient data was then compared to a synthetic control which was generated from at least two good quality control samples. If reference probe peaks ratios were outside of the normal range then the sample was failed, as was the case if there were high QC fragment peaks.<sup>263</sup>

Figure 2.1a shows an electropherogram with custom probes for *NCSTN* exons (on the left hand side of the trace) against the reference probes in the MRC-Holland P200 kit (on the right). Figure 2.1b shows the analysed data of this sample. The green dots represent the exons of the genes of interest (in this case *PSEN1* and *PSENEN*) and the blue dots represent the reference probes. As can be seen, there is always some slight variability in the peak height ratio of each exon but in this example they all fall within normal limits (represented by the horizontal green lines, MLPA peak height ratios defined as deletion  $<0.71$ , duplication  $> 1.23$ )<sup>263</sup>.



**Figure 2.1. Example of MLPA results.** A) A graph showing the raw peak heights (relative fluorescent units) relating to 5 exons of NCSTN (green) and reference probes (blue) in one individual. B) Graph showing the normalised peak ratios of each exon (control probes shown in blue). The normal limits are delineated by horizontal green lines. Duplications would be expected to be above the top line and deletions below the lower line.

#### 2.4.18 Preparation of Luria Bertani (LB) medium and 2 x LB media

Luria Bertani medium was prepared by adding 10g bacto-tryptone (Becton, Dickinson and Company, Sparks, USA), 5g yeast extract (Becton, Dickinson and Company) and 10g sodium chloride (VWR, Lutterworth, UK) to 800ml distilled water. pH was adjusted to 7.5 using sodium hydroxide (BDH Laboratory Supplies, Poole, UK) and the

final volume of media was made up to 1l with distilled water. The media was sterilised in an autoclave (200/35L, Boxer Laboratory Equipment Ltd., London, UK). 2 x LB medium was prepared by adding 20g bacto-tryptone, 10g yeast extract and 20g sodium chloride to 800ml distilled water. The pH was adjusted to 7.5, the volume made up to 1l and the media was sterilised as per 1x media.

#### **2.4.19 Preparation of C100-Flag pET-21b (-) plasmid**

A C100-Flag pET-21b (-) plasmid was provided by Dr Oliver Holmes, Centre for Neurological Diseases, Harvard Medical School, Boston, USA<sup>264</sup>. This plasmid was formed by generating a DNA fragment encoding amino acids 596-695 of the 695 amino acid isoform of amyloid precursor protein (APP), adding a Flag sequence (DYKDDDDK) at the C-terminus and inserting the whole fragment between Hind III and Nde I in to the prokaryotic expression vector pET-21b (Novagen, Merck Serono Ltd., Feltham, UK)<sup>265</sup>. The Methionine at the N-terminus (amino acid 596) serves as the translation start site. This plasmid was shipped on filter paper and extracted by placing the filter paper in 100ul TE buffer for 24 hours (4°C).

#### **2.4.20 Transformation of NEB 5-alpha competent E Coli with C100-Flag pET-21b (-) plasmid**

Bacterial transformation was performed by taking 50ul competent cells (NEB 5-alpha competent E Coli, New England Biolabs, C2988), adding 5ul C100F plasmid DNA and incubating on ice for 30s. The cells were heat-shocked at 42°C for exactly 45 seconds (to permeabilise the cell membranes and allow plasmid uptake) and incubated on ice. Cells were placed in 250ul SOC media and shaken at 37°C for 1 hour. 100ul was spread over a pre-warmed LB plate (25ml 1xLB + 25ul 100mg/ul Ampicillin (Sigma) spread over the surface). This plasmid confers resistance to ampicillin thus the use of an ampicillin containing plate encourages growth of the plasmid-expressing bacteria. When the plate was dry it was incubated at 37°C overnight (inverted) to allow the bacteria to

multiply. One colony was isolated on a pipette tip and added to 5ml LB. This was shaken (300rpm) at 37°C for 8 hours. A larger 150ml LB culture was inoculated with the whole 5 ml starter culture and shaken (300rpm) at 37°C for 16 hours. Bacterial cells were harvested by spinning at 6000 x g for 15 minutes at 4°C. The supernatant was discarded and plasmid was purified from the cell pellet as detailed below.

#### **2.4.21 Plasmid purification**

Plasmid purification was performed using the Qiagen HiSpeed Plasmid Midi Kit (Qiagen) (all reagents, spin columns and tubes stated in the below method are provided with this kit unless otherwise stated). The bacterial cell pellet was suspended in 6ml Buffer P1 and 6ml Buffer P2 was added. The tube was inverted and incubated at room temperature for 5 minutes to induce bacterial cell lysis. 6ml chilled Buffer P3 was added to the lysate and the sample was inverted 4 times to mix and then placed in a QIAfilter cartridge. The cartridge was incubated for 10 minutes at room temperature and filtered in to a HiSpeed Midi tip (removing precipitated protein, genomic DNA and detergent). The sample was allowed to enter the HiSpeed Midi tip resin by gravity flow (plasmid DNA is held within this resin). Plasmid DNA was washed by adding 20ml Buffer QC to the tip and then eluted with 5ml Buffer QF. DNA was precipitated by adding 3.5ml isopropanol (room temperature), mixing and incubating the sample for 5 minutes. This sample was added to a 20ml syringe and filtered through a QIAprecipitator Midi Module. The DNA was washed with 70% ethanol, dried (by pushing air through the QIAprecipitator) and eluted with 400µl TE Buffer. The purified plasmid was stored at -80°C.

#### **2.4.22 Transformation of E-Coli BL21 DE3 cells with C100Flag – pET-21b plasmid**

BL21 (DE3) competent E. coli (New England Biolabs, C2527I) were defrosted on ice and 100ng plasmid was added to 50µl of cells. The sample was incubated on ice for 30

minutes and then heat-shocked at 42 °C for exactly 10 seconds to permeabilise the cell membranes. The sample was placed in 950µl SOC media and shaken for 1 hour at 37°C. 100µl sample was spread over a pre-warmed LB plate (25µl 100mg/µl Ampicillin spread over the surface) and the plate was left inverted at 37°C overnight. A 5ml LB culture (+ 100µg/ml ampicillin) was inoculated with one island of bacterial cells and shaken at 180rpm for 8 hours at 37°C. 25µl of 100mg/ml ampicillin was added to this culture and it was stored at 4°C overnight. 4ml of this starter culture was warmed and added to 200ml warm 2 x LB media (+100µg/ml ampicillin). The culture was shaken at 180rpm at 37 °C until the OD<sub>600</sub> was 0.8 (3-4 hours, light absorbance measured on FLUOstar Omega plate reader, BMG Labtech, Almendgruen, Germany). Protein expression was induced by adding 1.0mM IPTG (SIGMA) to the culture and leaving it shaking at 180rpm at 37°C for 4 hours. The sample was spun at 3000 x g for 10 minutes to pellet the cells. The supernatant was discarded and the pellet was stored at -80°C.

## **2.5 Immunohistochemistry**

### **2.5.1 Embedding of skin biopsies in paraffin**

Skin biopsies were placed in formalin for 48 hrs and transferred to 70% EtOH. The sample was then processed in the Leica TP2020 Processor (standard machine protocol) and embedded in paraffin using the Leica EG1150C embedding machine.

### **2.5.2 Sectioning and staining paraffin embedded skin**

Paraffin sections were mounted on Superfrost Plus slides and left at 60°C overnight. The sections were de-waxed with xylene (2 x 5 minutes, VWR International) and 100% industrial methylated spirit (4 x 2 minutes, VWR International) and incubated in 3% hydrogen peroxide (VWR International) for 10 minutes (to remove any endogenous peroxidases). The slides were rinsed thoroughly with tap water. Slides were placed in boiling antigen retrieval working solution (2.1g citric acid in 1000ml dH<sub>2</sub>O) for 5 minutes and dried. Sections were covered with blocking solution for 5 minutes and then incubated in primary antibody overnight at 21°C. This was rinsed off with TBS and the slides were washed with TBS for 10 minutes. The sections were incubated in biotinylated goat anti-rabbit IgG secondary antibody (1:200, Vector Laboratories) diluted in blocking buffer (2g bovine serum albumin (Sigma), 20ml 0.5M TBS (pH7.6), 180 dH<sub>2</sub>O, 2ml 10% sodium azide) for 1 hour at room temperature. Slides were rinsed and washed in TBS for a further 10 minutes, then incubated in StreptABComplex-HRP (10µl Reagent A (Streptavidin), 10µl Reagent B (Biotinylated HRP), 1ml TBS) (Vectorlabs) for 30 minutes at room temperature. The slides were rinsed and washed with TBS for 5 minutes and then developed by being immersed in DAB solution (1ml 5% DAB (VWR International), 200ml Tris buffer (0.1M, pH6), 200µl 100vol H<sub>2</sub>O<sub>2</sub>) for 10 minutes (with gentle agitation). After a 5 minute wash with water, sections were counterstained with haematoxylin for 2 minutes and washed until clear. Sections were dehydrated by rinsing with 100% IMS (4 x 2 minutes), rinsed in xylene (2 x 5 minutes)

and mounted with distyrene plasticizer and xylene (DPX, SIGMA). Antigen staining appears brown and nuclei appear blue.



## **2.6 Cell Culture**

### **2.6.1 Isolation of primary fibroblasts from skin biopsies**

Skin samples were harvested and placed in PBS. They were transferred to an uncoated 10cm petri dish and washed with PBS (Life Technologies). Subcutaneous fat and any loose connective tissue was detached with a scalpel and tweezers leaving epidermis and dermis. The skin was placed in 10ml Dispase 2 (Stemcell Technologies, Grenoble, France) overnight at 4°C. The epidermis was peeled from the dermis with tweezers. The dermis was cut in to small fragments, dried and distributed over the coated surface of a scored 25cm<sup>2</sup> tissue culture flask (Corning). After 30 seconds, Cascade Biologics media 106 (Gibco) + LSGS supplement kit (Life Technologies, this media always has this supplement kit added unless otherwise stated) was slowly added to the flask (so as not to disturb the skin fragments) and the flask was placed in a 37°C (5% CO<sub>2</sub>) incubator. After 24 hours the lid on the flask was slightly unscrewed to allow the cells to breathe. Media was replaced after seven days and every three days thereafter. When sufficient cell numbers had grown out from the skin fragments, cells were detached and passaged (section 2.6.2).

### **2.6.2 Passage of primary human fibroblasts**

Cell media was aspirated and the cells were washed with 5-10ml sterile PBS (Gibco, Life Technologies). The cells were then incubated in 5-7ml TrypLE Express 1x (Gibco, Life Technologies) for 5 minutes to induce cell detachment. Once the cells had detached they were mixed with an equal volume of DMEM (1x) + GlutaMAX supplemented with 10% FBS (Biosera, Uckfield, UK), penicillin (250 units/ml) and streptomycin (250µg/ml) (this media always contains these supplements unless otherwise stated) to inactivate the TrypLE Express. The cells were transferred to a 15ml centrifuge tube and centrifuged at 1000rpm for 5 minutes at room temperature. The supernatant was aspirated and the cell pellet was suspended in 5 ml Cascade Biologics medium 106

(Gibco, Life Technologies) and re-seeded (Corning tissue culture flasks) or frozen in liquid nitrogen for storage as detailed in section 2.6.5.

### **2.6.3 Maintenance of primary human fibroblasts**

Cells were maintained in Cascade Biologics medium 106 (Gibco, Life Technologies) supplemented with EGF, hydrocortisone, insulin, FCS, gentamycin and amphotericin B (LSGS kit S-003-K, Gibco, Life Technology) (when this media is mentioned in the remainder of the methods section it always contains these supplements unless otherwise stated). Cell media was aspirated and changed every 3 days. Tissue culture flasks were kept at 37°C (5% CO<sub>2</sub>). Cells were passaged when they became >80% confluent. Flasks were discarded if any sign of infection was observed.

### **2.6.4 Maintenance of S20 cell line**

S-20 is a Chinese hamster ovary (CHO) cell line that stably expresses Nicastrin-V5/His, human PSEN1, N-terminally FLAG-tagged Pen-2 and C-terminally HA-tagged Aph-1 $\alpha$ <sup>266</sup>. As it over-expresses the gamma-secretase components it can be used as a positive control in gamma-secretase activity assays. Cells were maintained in DMEM (1x) + GlutaMAX. Cell media was aspirated and changed every 2 days. S1 and S20 cells were otherwise maintained and passaged as stated for primary human fibroblasts.

### **2.6.5 Storing cells in liquid nitrogen**

Cells were detached, pelleted and suspended in FBS (Biosera)/ DMSO (Sigma) (90%FBS/ 10%DMSO). 1ml aliquots (1ml cryo-vials, Corning) were frozen at -80°C in Mr Frosty Freezing containers (Thermo Scientific) containing 100% isopropyl alcohol (Sigma). Cells were transferred to liquid nitrogen for long term storage. Cells were recovered by briefly warming at 37°C, suspending in 15ml DMEM (to wash DMSO off) and spinning at 1000rpm for 5 minutes at room temperature. The supernatant was

aspirated and the cell pellet suspended in 15ml growing media and seeded in a tissue culture flask (Corning).

## **2.7 Protein analysis**

### **2.7.1 Isolation of total protein from primary human fibroblasts in culture**

Cell media was aspirated and the cells were washed with PBS (Life Technologies). Cell detachment was induced with TrypLE Express 1x (Gibco). Once the cells had detached they were mixed with an equal volume of medium containing serum (DMEM (1x) + GlutaMAX) to inactivate the TrypLE Express. The cells were transferred to a 15ml centrifuge tube and centrifuged at 1000rpm for 5 minutes at room temperature. The supernatant was aspirated and the cell pellet was suspended in 100-150µl RIPA buffer (Sigma) for 20 minutes on ice. The sample was transferred to a 1.5ml micro-centrifuge tube, spun at 15000rpm for 30 minutes at 4°C and frozen at -80°C.

### **2.7.2 Isolation and solubilisation of cell membrane proteins (including gamma-secretase microsomes) from cells in culture**

Primary human fibroblasts were cultured and passaged until the predicted cell count reached around  $5 \times 10^7$  cells for each individual (roughly 18 x 175cm<sup>2</sup> tissue culture flasks, Corning). Based on previous publications, this cell number was predicted to harvest a quantity of gamma-secretase sufficient for enzyme activity assays<sup>264</sup>. Cell media was aspirated from each flask and cells were washed with PBS (Life Technologies). Cell detachment was induced with TrypLE Express 1x (Gibco). Upon cell detachment an equal volume of medium containing serum (DMEM (1x) + GlutaMAX) was added to inactivate the TrypLE Express. The cells from each flask were transferred and combined in 50ml centrifuge tubes and centrifuged at 250g for 10 minutes at room temperature. The supernatant was aspirated and cells were re-suspended in 2.5ml ice cold MES buffer (50mM MES pH 6 (Sigma), 150mM NaCl (VWR), 5mM MgCl<sub>2</sub> (Sigma), 5mM CaCl<sub>2</sub> (Sigma) with complete protease inhibitor cocktail with EDTA (Roche)). Cell suspensions were combined in one 50ml tube and kept on ice. Cells were lysed using a 15ml glass homogeniser (>30 plunges). The nuclei

and unbroken cells were harvested by centrifuging the sample at 3000g for 10 minutes at 4°C. The postnuclear supernatant was centrifuged at 100000g for 1 hour (4°C) to pellet the total cell membranes (L8-60M Ultracentrifuge, Beckman Coulter Ltd, High Wycombe, UK). The supernatant was discarded and the pellet was fully re-suspended in 5ml ice cold sodium bicarbonate buffer (0.1M NaHCO<sub>3</sub> pH 11.3, Sigma) and incubated at 4°C for 20 minutes. The washed membranes were pelleted by spinning at 100000g for 1 hour (4°C) and fully re-suspended in 140µl solubilisation buffer (1% CHAPSO (Sigma), 50mM HEPES pH7 (Sigma), 150mM NaCl (VWR), 5mM MgCl<sub>2</sub> (Sigma), 5mM CaCl<sub>2</sub> (Sigma)). The sample was incubated for 1 hour at 4°C. Insoluble material was pelleted by spinning the sample at 100000g for 1 hour (4°C) and the supernatant (soluble fraction) was aliquoted, snap frozen on dry ice and stored at -80°C. For the purpose of performing gamma-secretase enzyme assays this supernatant is defined as soluble gamma-secretase preparation (used as crude gamma-secretase fractions in enzyme activity assays).

### **2.7.3 Protein quantification**

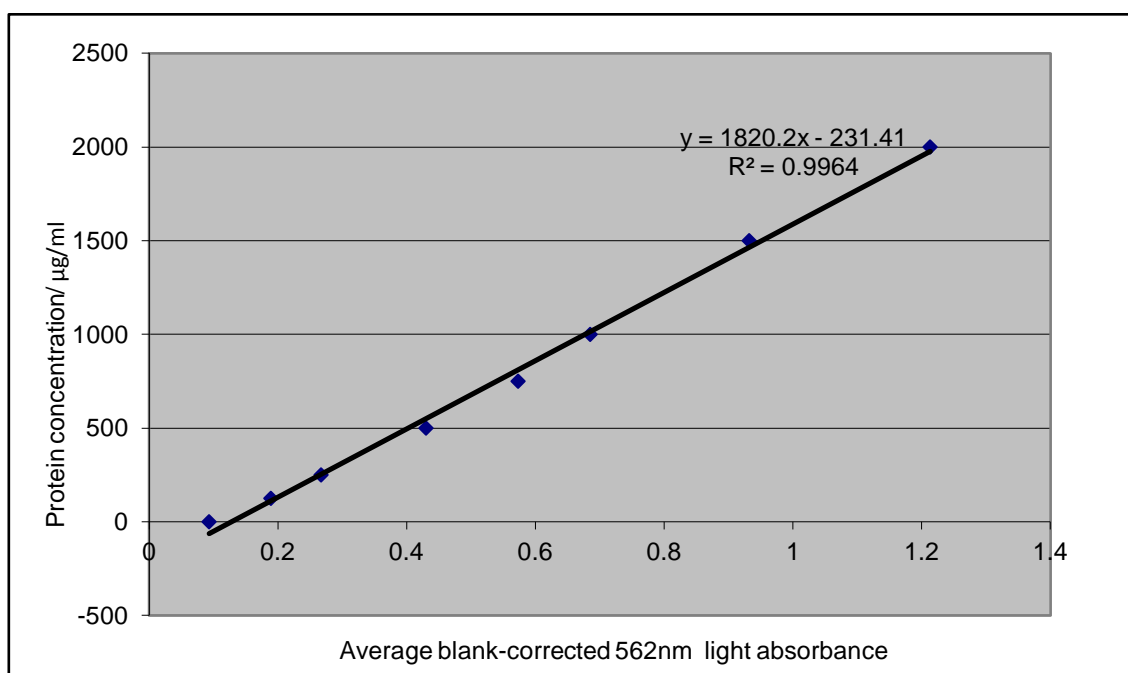
Protein samples were quantified using the Pierce Microplate Bicinchoninic Acid (BCA) Protein Assay Kit (Thermo Scientific, Loughborough, UK). BCA working reagent was made up by mixing 50 parts reagent A with 1 part reagent B. 200 µl BCA working reagent was mixed with 10µl of each protein sample in a 96 well plate (samples mixed well by pipetting up and down). The protein samples in every run comprised 7 pre-made albumin standards (Thermo Scientific), the protein samples to be quantified and a control sample (10ul BCA working reagent). All runs were performed in duplicate. The plate was incubated at 37°C for 30 minutes and cooled at room temperature for 5 minutes.

The absorbance of the standards, protein samples to be quantified and the control sample was measured at 562nm on a plate reader. The two readings generated for the

duplicate runs performed for each sample were then averaged. If the two readings differed significantly for any one sample the experiment was repeated for that sample. Table 2.10 shows an example of the light absorbance readings achieved with the protein standards. The standard control was used as a blank. The average 562nm absorbance value of the control sample was subtracted from the average 562nm absorbance value of each of the standards and the protein samples to be quantified. A standard curve was drawn by plotting the average blank-corrected 562nm value for each albumin standard versus its concentration ( $\mu\text{g/ml}$ ) (Figure 2.2). The standard curve was used to determine the protein concentration of each unknown sample.

**Table 2.10. Table demonstrating variation in light absorbance with protein standards**

Sample	Protein Concentration ( $\mu\text{g/ml}$ )	Average absorbance at 562nm
Blank	0	0.093
Standard 1	125	0.189
Standard 2	250	0.267
Standard 3	500	0.43
Standard 4	750	0.573
Standard 5	1000	0.685
Standard 6	1500	0.932
Standard 7	2000	1.213



**Figure 2.2. Example of a standard curve plotted for albumin protein standards.** The curve was drawn by plotting the average blank-corrected 562nm value for each protein standard versus its concentration (µg/ml).

#### 2.7.4 Immunoblotting Nicastrin and PEN-2

Protein samples were thawed on ice and the desired concentration/ quantity of protein (20µg for NCSTN, 25µg for PEN-2) was mixed with 6x Laemmli buffer (Sigma) (5 parts protein to 1 part 6x Laemmli buffer) and heated to 95 degrees for 5 minutes. This denatured the protein in the sample, removing secondary and tertiary protein structure, thus allowing proteins to be separated according to their molecular weight. Negatively charged Sodium dodecyl sulphate within the Laemmli buffer attached to the sample protein thus enabling separation of the proteins by electrophoresis. Samples were quickly spun down and loaded in to the wells of a 10% Tris-HCL Ready Gel (Biorad, Hemel Hempstead, UK) for NCSTN and a 4-20% gradient Ready Gel (Biorad) for PEN-2. The first and last well of each gel were loaded with 7 µl of pre-stained protein ladder (Protein Plus Protein Standards Dual Colour, Biorad) which enabled the detection of proteins between 10 and 250 kDa. Gels were placed in a running tank

(Biorad) with 1x Tris/ Glycine/ SDS buffer (Biorad) and an electronic current was passed through the gel (100V for 5 minutes, increased to 200V for the remaining time) to separate the proteins by molecular weight. Protein was transferred onto an ECL membrane using a Trans-blot TURBO Transfer System Transfer Pack (Biorad). The membrane was washed in distilled water for 10 minutes and then blocked in 5% non-fat dry milk (Marvel original, Marvel, Spalding, UK) in TBS-T for 30 minutes at room temperature. The membrane was washed in TBS-T for 30 minutes (3 x 10 minute washes) at room temperature and incubated in primary antibody in 5% non-fat dry milk in TBS-T overnight at 4°C (NCSTN Ab24741 (Abcam), PSENEN Ab18189 (Abcam)). The membrane was washed in TBS-T for a further 30 minutes (3 x 10 minutes) and incubated in a horseradish peroxidase-linked anti-rabbit secondary antibody (NA934V, GE Healthcare UK Ltd) for 1 hour at room temperature. The membrane was washed in TBS-T for 30 minutes (3 x 10 minutes) and protein bands were visualised using ECL Prime Western Blotting Reagent (GE Healthcare).

### **2.7.5 Immunoblotting $\beta$ -actin**

Following the immunoblotting of NCSTN and PEN-2, the same membranes were blotted for  $\beta$ -actin.  $\beta$ -actin is endogenously expressed and therefore acts as a loading control, reflecting the total amount of protein originally loaded for each sample. This information is required to calculate any inter-individual variation in NCSTN and PEN-2 expression. Following the visualisation of NCSTN and PEN-2 bands, membranes were washed in TBS-T for 10 minutes. They were then placed in stripping buffer (30ml Tris-HCL (Sigma) pH 6.8, 3.5ml  $\beta$ -mercaptoethanol (Sigma), 50ml 20% SDS (Sigma), 416.5ml dH<sub>2</sub>O) at 50 °C for 30 minutes. The membrane was washed for a further 10 minutes in TBS-T and then incubated in primary antibody (1:1000, 4967S, Cell signalling Technology) in 5% non-fat dry milk in TBS-T for 1 hour at room



temperature. The remaining process was as described for day two of immunoblotting NCSTN and PEN-2 (section 2.7.4).

#### **2.7.6 Immunoblotting Flag-tagged proteins**

As described for the immunoblotting of NCSTN and PEN-2 with the following variations. Protein samples were run on a 4-20% gradient Tris-HCL Ready Gel (Biorad). The ECL membrane was incubated in primary antibody (F3165, Sigma) in 5% non-fat dry milk in TBS-T for 1 hour at room temperature.

#### **2.7.7 Analysis of western blots**

Band density was measured using Image J software (National Institute of Health, USA). All analyses were performed in triplicate and an average value was calculated for each band. The density of the protein band in question (ie. NCSTN or PEN-2) divided by the density of the loading control ( $\beta$ -actin) band represented to the total amount of protein present.

## 2.8 Cell based assays

### 2.8.1 Addition of cycloheximide to assess for nonsense mediated decay

1 000 000 primary human dermal fibroblasts (suspended in Cascade Biologics media 106, Gibco, Life Sciences) were plated in 10cm diameter circular tissue culture dishes and left overnight to facilitate attachment. Plates were labelled as time = 0, 6 and 24 hours. Cell media was aspirated from all three dishes. 10ml of Cascade Biologics media 106 was replaced in the T=0 plate. 10ml Cascade Biologics media 106 plus 20µg/ml cycloheximide (Sigma) was added to each of the remaining two plates. RNA was isolated from the t = 6 hours cells after 6 hours and the t = 0 and 24 hours cells at 24 hours as previously described in section 2.4.11. Total *NCSTN* and *PSENEN* mRNA levels were assessed by multiplex real-time PCR as previously described (section 2.4.16). Cycloheximide is an inhibitor of protein biosynthesis that acts by inhibiting the translocation stage of protein synthesis. As a consequence it is known to inhibit nonsense-mediated decay<sup>267</sup>. If nonsense mediated decay is present then the expression of the relevant gene transcript should incrementally increase the longer the cells are exposed to cycloheximide (up to 24 hours).

## **2.9 Gamma-secretase enzyme activity assays**

### **2.9.1 Purification of C100-Flag substrate**

The cell pellet described in section 2.4.22 was re-suspended in 10ml lysis buffer (250µl 1M Tris pH 7.0 (Sigma), 750µl 5M NaCl (VWR), 250µl Triton X-100 (Sigma), 250µl bacterial protease inhibitors (P8465 5ml protease inhibitors, SIGMA), 23.5ml dH<sub>2</sub>O) and homogenised with a 15ml glass homogeniser. The sample was spun at 3000 x g for 10 minutes. The cell lysate supernatant (containing protein) was removed and stored at 4°C and the pellet (insoluble fraction) was discarded. 4ml M2 agarose slurry (A2220, SIGMA) was washed with 10ml lysis buffer (4°C) and spun at 1000 x g for 5 minutes. The supernatant was removed leaving 2ml equilibrated M2 agarose beads. The cell lysate supernatant was added to the 2ml equilibrated M2 agarose beads and mixed at room temperature for 1-2 hours (allowing the C100-Flag to bind to the beads). The sample was transferred to a drip column (BioRad 731-1550, Biorad) and the unbound fraction was allowed to drip through (discarded). The drip column membrane was washed 5 times with 5ml lysis buffer and the C100-Flag protein was eluted five times with five separate 2ml volumes of elution buffer (110µl Igepal (Sigma), 1.1ml 1M glycine pH 2.7 (Sigma), 9.8ml dH<sub>2</sub>O). Samples were stored at -80°C. The agarose beads were washed with 10ml PBS, suspended in storage buffer (5ml Glycerol (Sigma), 1ml 10 x PBS (Life Technologies), 4ml dH<sub>2</sub>O, 20µl Sodium Azide (Sigma)) and kept at -20 °C (could be re-used once). Immunoblotting of the elutes was performed to assess for the presence of purified C100-Flag (as per anti-Flag western protocol section 2.7.6). C100-Flag would be used as a substrate in gamma-secretase activity assays.

### **2.9.2 Gamma-secretase activity assay**

6µl of L- $\alpha$ -phosphatidylethanolamine chloroform (PC, 10mg/ml) and 1.5µl L- $\alpha$ -phosphatidylethanolamine chloroform (PE, 10mg/ml) (Avanti Polar Lipids, Alabaster, USA) were dried under Nitrogen for 10 minutes. Each lipid was dissolved in 6µl and

3µl of HEPES buffer (50mM HEPES pH 7.4, 150mM NaCl) respectively, incubated at room temperature for 10 minutes and sonicated (XB3 Ultrasonic Bath, Grant, Shepreth, UK) at room temperature for 10 minutes. Cholesterol (C4951, SIGMA) was dissolved to a concentration of 31.75mg/ml in water. The protein concentration of the soluble gamma-secretase preparation (described in section 2.7.2) obtained from each individual was calculated using a BSA assay. 8µg soluble gamma-secretase preparation was added to 5µl PC, 2.5µl PE and 2µl cholesterol and the mix was made up to 50ul with HEPES buffer.

Final reaction concentrations:

PC 1mg/ml

PE 0.25mg/ml

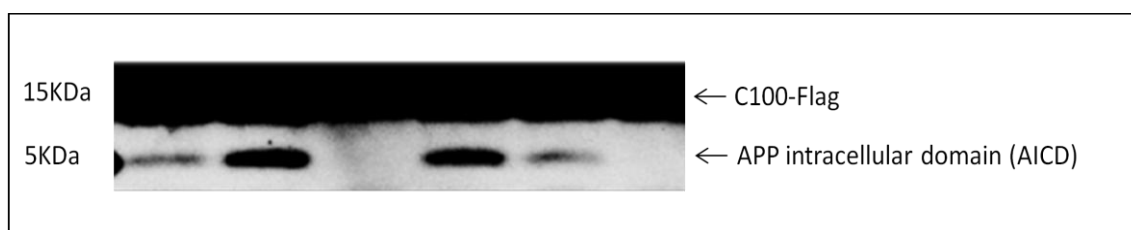
Cholesterol 0.065mg/ml

CHAPSO 0.25%

This mixture was allowed to incubate at 32°C for 30 minutes and on ice for a further 10 minutes. 1ul C100-Flag peptide (see section 2.9.1) was added to the mix to act as the enzyme substrate and the reaction was left to incubate at 37 °C for 4 hours. A 12µl aliquot of reaction was mixed with 12µl 2 x Laemmli buffer and frozen at -80°C. All samples for analysis were run in the same assay, using the same reagents. Experiments were performed four times on separate occasions to ensure consistency. Soluble gamma-secretase preparation obtained from S20 cells (a cell line known to over-express the gamma-secretase components) was run with every assay as a positive control. A gamma-secretase inhibitor (L458) was added to some reactions as a negative control.

Immunoblotting of the reaction mixtures (mixture + Laemmli buffer totalling 24µl) was performed using an anti-FLAG antibody (F3165, Sigma) as detailed in section 2.7.6. Enzyme activity assays were ultimately performed by collaborators in the Centre for

Neurological Diseases, Harvard Medical School, USA, as I was unable to successfully generate enzyme activity data in-house. They developed immunoblots using a Li-Cor Odyssey system (Li-Cor, Lincoln, USA), a representative blot is shown in Figure 2.3. The large band detectable at 15kDa is the C100-Flag substrate. The smaller band viable at 8kDa is an enzyme product, amyloid precursor protein intracellular domain (AICD). When gamma-secretase cleaves APP ( $\epsilon$ -cleavage) it produces AICD and Amyloid- $\beta$  peptides. The amount of AICD produced in the assay is regarded as a measure of total gamma-secretase activity.



**Figure 2.3. An example western blot generated from a gamma-secretase activity assay.** The C100-Flag substrate can be seen at 15kDa and the AICD product at 5 kDa.

Amyloid- $\beta$  peptides are sequentially shortened (processed) by gamma-secretase (carboxy-peptidase activity). Enzyme-linked immunosorbent assays (ELISA) were performed by our collaborators at Harvard Medical School, USA, to quantify the levels of two such peptides, A $\beta$ 40 and A $\beta$ 42, in the reaction mixtures (Human Abeta40 and Abeta42 ELISA kits, Life Technologies).

## **2.10 Statistical analysis**

Bar graphs were drawn and error calculations were performed using Microsoft Excel. Statistical analysis was performed using GraphPad Prism 6 (GraphPad Software Inc, La Jolla, USA). All real-time PCR experiments were performed in triplicate and repeated three times. The average fold change and standard deviation across all three repeat experiments was calculated. Unpaired t-tests were performed to assess the statistical significance of the results generated. In-house immunoblots were analysed using Image J software (NIH, USA). Band densities were recorded three times from which average band densities and standard deviation were calculated. Statistical significance was assessed using unpaired t-tests. Immunoblots generated for the purpose of assessing gamma-secretase enzyme activity were analysed by collaborators in the Centre for Neurological Diseases, Harvard Medical School, using the Li-Cor Odyssey system analysis software (Li-Cor). All data analysis was performed in house. Average values and standard errors were calculated using data from all seven repeat enzyme activity assays. Statistical analyses were performed using one-way ANOVA tests (Kruskal-Wallis with Dunn's correction).

## **Chapter 3: Results: Mutations in the gamma-secretase genes *PSENEN* and *NCSTN* underlie some familial forms of hidradenitis suppurativa**

### **3.1 Introduction**

HS is inherited in an autosomal dominant manner in some families<sup>132</sup>. Molecular genetic analysis of such families always therefore had the potential to further our understanding of disease pathogenesis. The benefits of this approach were realised in 2010 when heterozygous mutations were reported in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1* in six Chinese multiplex kindreds with HS<sup>135</sup>. All six kindreds demonstrated full co-segregation. In parallel with the publication of that data we had recruited seven extended British Caucasian pedigrees and were planning experiments to identify novel disease-causing genes.

#### **3.1.1 Hypothesis and objectives**

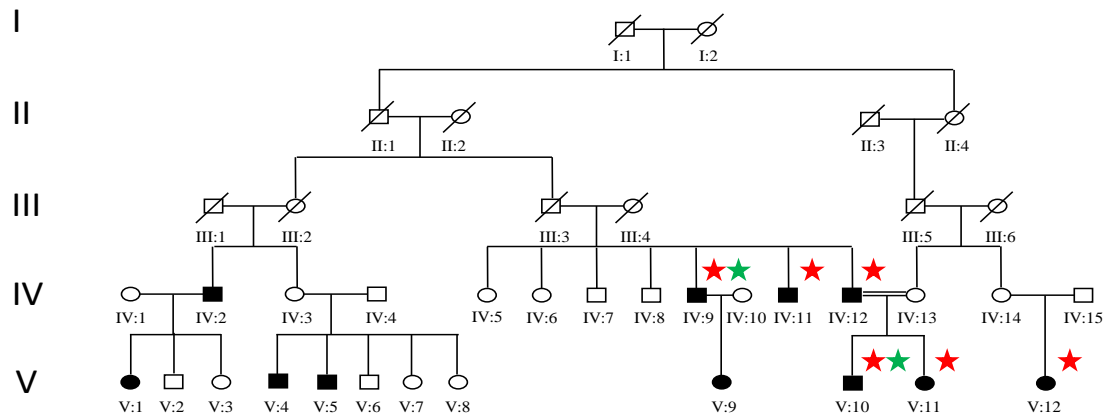
In light of the Chinese data, it was hypothesised that mutations in *NCSTN*, *PSENEN* and *PSEN1* underlie British familial cases of HS. The aim of this work was therefore to investigate the role of mutations in *NCSTN*, *PSEN1* and *PSENEN* in seven British kindreds with HS. As discussed in the introduction, gamma-secretase is a four sub-unit enzyme encoded by six different genes. It could therefore have been hypothesised that the remaining genes, *APH1A*, *APH1B* and *PSEN2*, also underlie some forms of HS. That hypothesis was not interrogated at that stage, partly due to the size of the genes involved.

## **3.2 Patient cohort and samples**

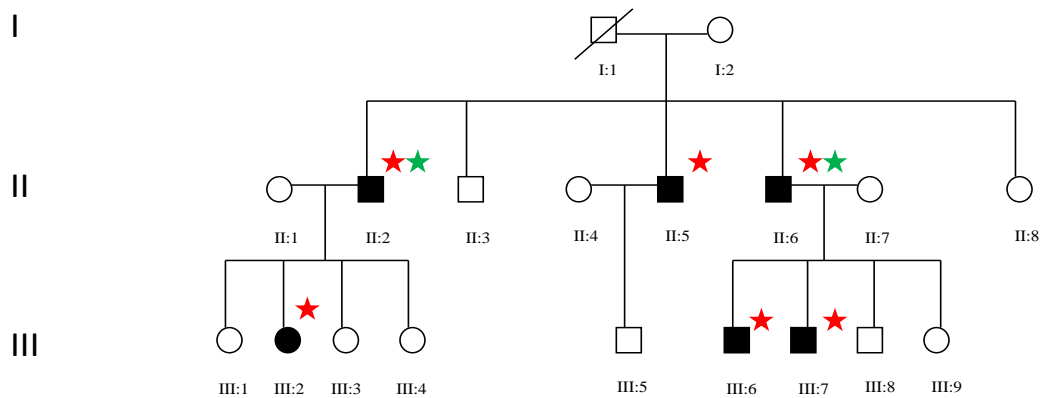
Fifty three individuals were recruited from seven multigenerational pedigrees in which HS was inherited as an autosomal dominant trait (see Figure 3.1, Figure 3.2, Figure 3.3). All individuals were of British origin and 27 of the 53 individuals met diagnostic criteria.



## Pedigree 1

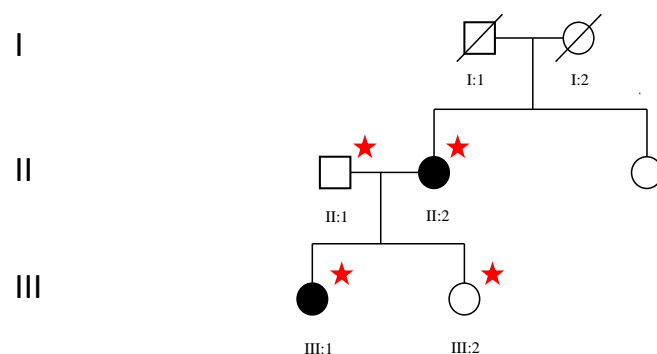


## Pedigree 2

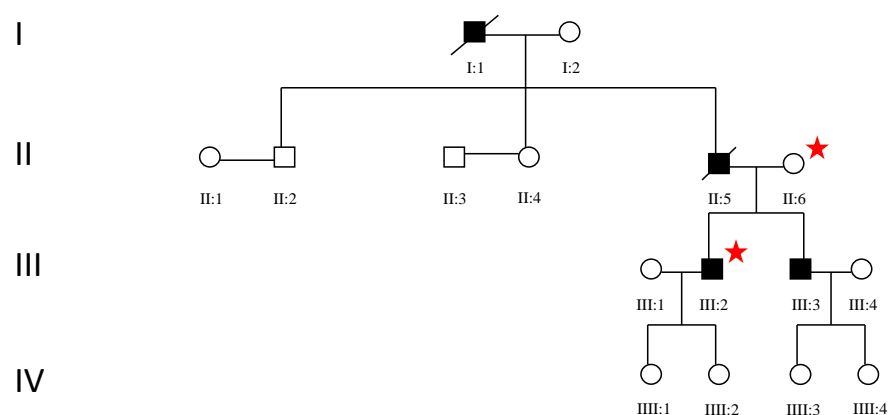


**Figure 3.1. Family pedigrees recruited for genetic analysis.** Red stars mark the individuals in whom *NCSTN*, *PSENEN* and *PSEN1* were sequenced. Green stars mark individuals assessed for large scale deletions and duplications in those genes via multiplex ligation-dependent probe amplification (MLPA).

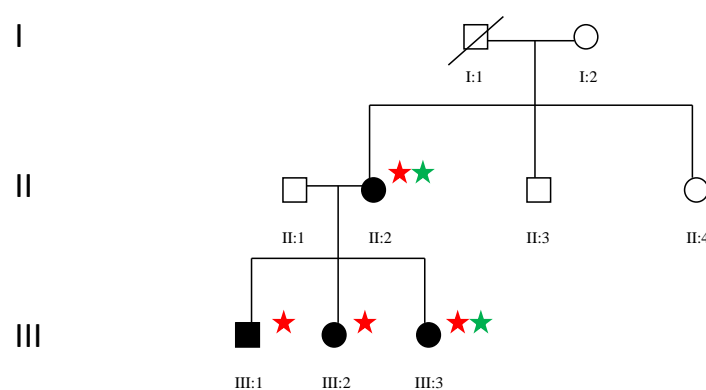
Pedigree 3



Pedigree 4

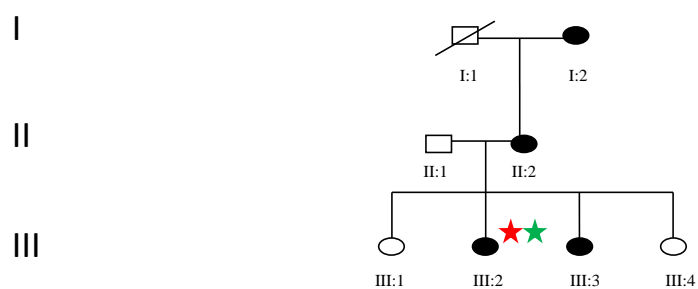


Pedigree 5

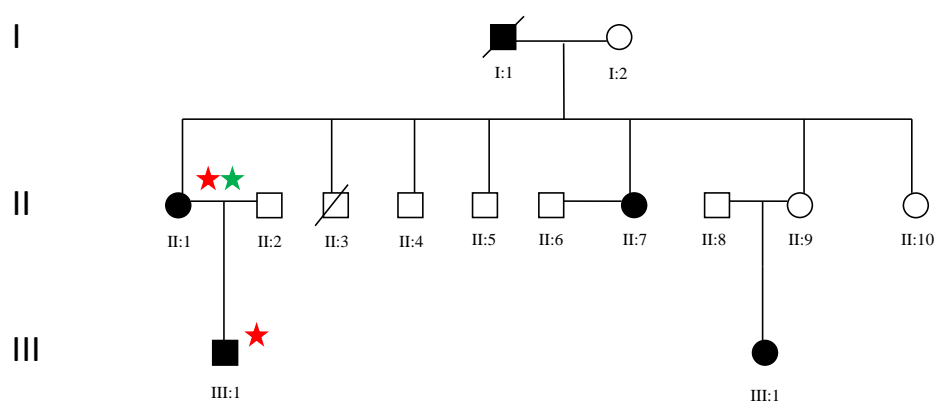


**Figure 3.2. Family pedigrees recruited for genetic analysis.** Red stars mark the individuals in whom *NCSTN*, *PSENEN* and *PSEN1* were sequenced. Green stars mark individuals assessed for large scale deletions and duplications in those genes via MLPA.

### Pedigree 6



### Pedigree 7



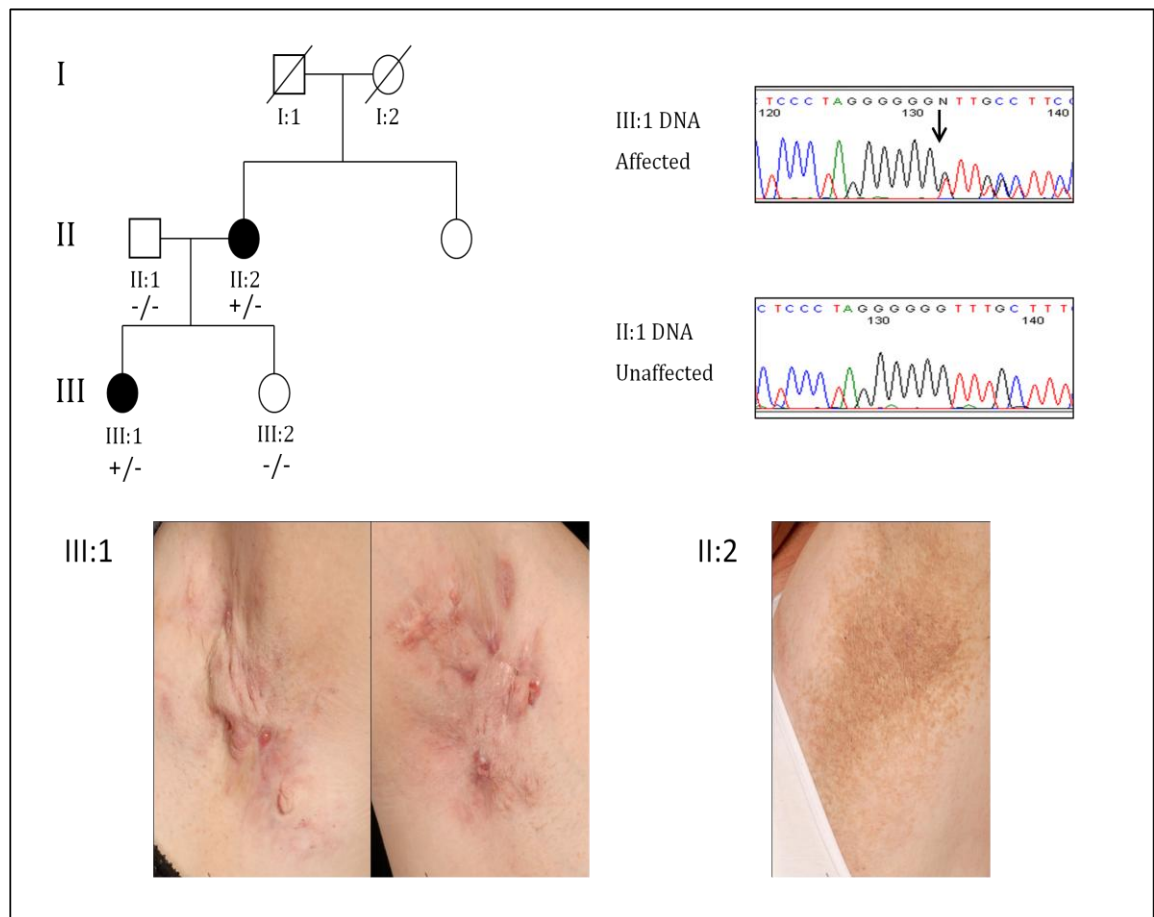
**Figure 3.3. Family pedigrees recruited for genetic analysis.** Red stars mark the individuals in whom *NCSTN*, *PSENEN* and *PSEN1* were sequenced. Green stars mark individuals assessed for large scale deletions and duplications in those genes via MLPA.

Study participants provided blood or saliva samples for DNA. All coding regions of *NCSTN* (17 exons), *PSEN1* (10 exons) and *PSENEN* (3 exons) were amplified by PCR (polymerase chain reaction) using exon flanking intronic primers and Sanger sequenced in 19 affected individuals, one or more from each of the seven pedigrees (red stars in Figure 3.1, Figure 3.2 and Figure 3.3 indicate individuals in whom the genes were sequenced). Sequencing revealed a novel heterozygous single nucleotide insertion in *PSENEN* (c.66\_67insG) and a novel single nucleotide substitution in the *NCSTN* exon 9/intron9 donor splice site (c.1125+1 G>A) in each of two pedigrees. No other novel or rare variants (population frequency of <1%, defined as variants with a frequency < 1% in both the 1000 Genomes Project database<sup>260</sup> and the NHLBI Exome Sequencing Project Exome Variant Server (EVS, <http://evs.gs.washington.edu/EVS/>)) were detected in coding regions or within 10 base pairs of splice boundaries in those genes in the remaining five pedigrees (pedigrees 1, 2, 5, 6, and 7).

### 3.3 Identification of a novel heterozygous single base insertion in *PSENEN* (c.66\_67insG)

Individuals II:2 and III:1 from pedigree 3 (Figure 3.2) were both heterozygous for a novel insertion in *PSENEN* (c.66\_67insG, Figure 3.4). The insertion is predicted to lead to a frameshift and an altered protein product (p.Phe23ValfsX98). The mutation co-segregated with disease status in this kindred (Figure 3.4) and was absent in 200 control chromosomes of European ancestry.

Both affected individuals from pedigree 3 developed painful nodules, pustules, abscesses and resultant scars in their axillae, under their breasts and in their groins from the age of 15. III:1 was more severely affected, having additional involvement of her buttocks and sinus tracts in the axillae and groin. They had a Body Mass Index (BMI) of 33 (II:2) and 23 (III:1). A BMI of 18 – 25 is considered to be within the normal range, 25-30 overweight and >30 obese. Therefore individual II:2 would be classed as obese but her daughter (III:1) was within the normal range. Neither individual had ever smoked. II:2 felt that her symptoms worsened around menstruation, whereas her mother, III:2, noticed no cyclical change. On examination III:2 had an atypical speckled macular hyperpigmentation in her axillae and groin and II:2 had similar hyperpigmentation in the groin. Patient III:2 had failed to respond to multiple antibiotics (flucloxacillin, clindamycin and rifampicin), the oral contraceptive, isotretinoin and ciclosporin. She did however respond to treatment with the anti-TNF agent Infliximab and surgical excision of the affected areas in her axillae and groin was effective. Patient II:2 had previously been treated with multiple antibiotics to little effect but no longer required treatment as her disease was quiescent.

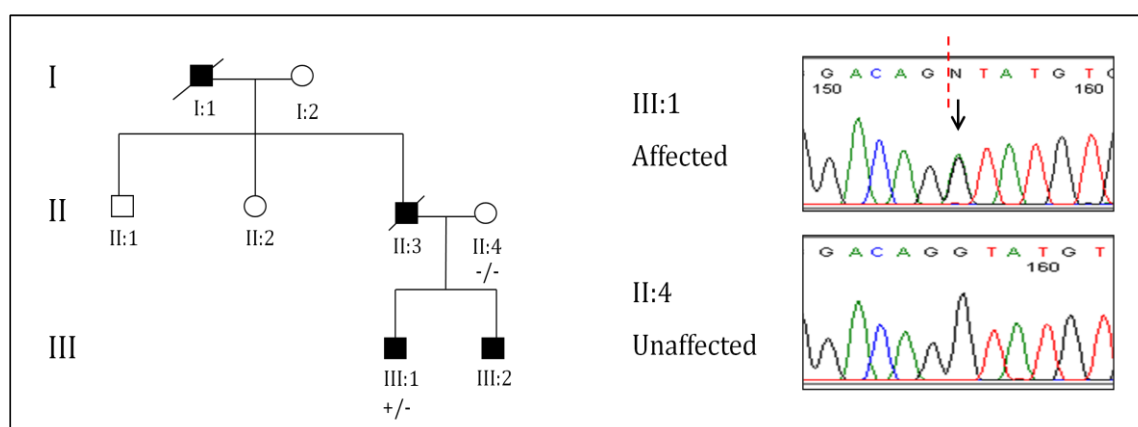


**Figure 3.4. The *PSENEN* c.66\_67insG mutation was detected in affected individuals from Pedigree 3.** Sequencing of *PSENEN* exon 3 revealed the c.66\_67insG variant (arrow) in genomic DNA from III:1 and II:2 and wild type sequence in unaffected individuals II:1 and III:2. Individual III:1 had inflamed nodules, abscesses, fistulae and scarring in the axillae. Individual II:2 had multiple open comedones and an inflammatory nodule on the background of a speckled, macular hyperpigmentation in the axillae (photos courtesy of Guy's and St. Thomas' NHS Foundation Trust, all individuals consented to publication).

### 3.4 Identification of a novel heterozygous single nucleotide substitution in the *NCSTN* exon 9/intron9 donor splice site (c.1125+1 G>A)

The proband from Pedigree 4 (Figure 3.2, III:1) was heterozygous for a novel single nucleotide substitution in *NCSTN* (c.1125+1 G>A). The substitution was of the first base of intron 9, in the highly conserved donor splice site. The mutation was absent in his unaffected mother (Figure 3.4, II:4) but no other members of the kindred were available for analysis to confirm co-segregation of the variant with disease status. The variant was not observed in 200 control chromosomes of European ancestry.

Individual III:1 had HS from the age of 16. He reported cysts, painful nodules, pustules, abscesses, sinus tracts and scarring over his neck, axillae, suprapubic area, groin, buttocks and thighs. He had a body mass index of 30 (overweight) and a 15 pack year smoking history. His disease had failed to respond to multiple antibiotics (flucloxacillin, erythromycin, oxytetracycline), acitretin and isotretinoin. He had not received immunosuppressant medication or anti-TNF $\alpha$  therapy.



**Figure 3.5.** The *NCSTN* c.1125+1 G>A mutation was detected in an affected individual from pedigree 4. Sequencing revealed a novel single nucleotide substitution (c.1125+1 G>A, see arrow) of the first base of the exon 9/ intron 9 donor splice site (the red dotted line marks the end of the exon) in affected individual III:1. No substitution can be seen in the only available unaffected individual (II:4).

### **3.5 No large scale deletions or duplications in *NCSTN*, *PSENEN* and *PSEN1* were detected in any of the seven kindreds**

Novel variants were detected in the coding regions of *NCSTN* and *PSENEN* in affected individuals from two of the seven multiplex kindreds. No variants (<1% population frequency) were detected in *PSEN1* in these families. The primers used for PCR and the standard Sanger sequencing method employed in this study were capable of detecting substitutions and small insertion/ deletions in coding regions and flanking splice sites. Large deletions and duplications of exons and genes would be missed using this screening technique. As a consequence, we hypothesised that whole gene or exon deletions and duplications in *NCSTN*, *PSENEN* and *PSEN1* underlie HS in the five kindreds in which no mutations were detected.

#### **3.5.1 Methodological approaches to detecting copy number variation (CNV)**

There are a number of experimental approaches that can be employed to detect copy number variation (CNV). These comprise fluorescence *in situ* hybridisation (FISH), array-comparative genomic hybridisation, single nucleotide polymorphism (SNP) arrays, multiplex ligation-dependent probe amplification (MLPA) and, more recently, next generation sequencing techniques (such as exome sequencing). The accuracy of CNV detection from exome sequencing data is dependent upon the read depth generated and the bioinformatic analyses that are employed. At the time of study, there was insufficient expertise and experience to consider that a viable option. MLPA was chosen as it is a fast, accurate and sensitive technique that can be used to look at multiple exons simultaneously<sup>268</sup>. The local clinical diagnostic laboratory routinely perform MLPA and consider it to be the most reliable and efficient method for detecting CNV<sup>263</sup>.

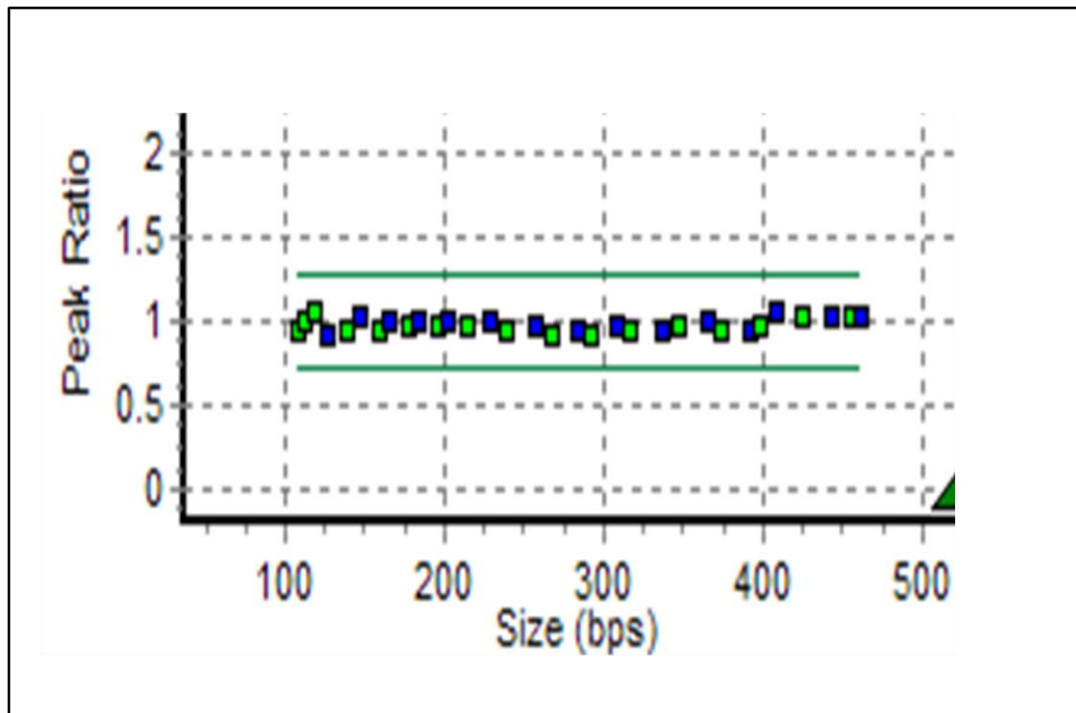


### **3.5.2 Multiplex ligation-dependent probe amplification (MLPA)**

MLPA is a multiplex-PCR based method of detecting CNV<sup>268</sup>. Probes are designed (forward and reverse) to target specific regions (in this case *NCSTN* exons 1-17, *PSENEN* exons 1-3 and *PSEN1* exons 1-10). These target probes are then run with reference probes which bind to known non-variable regions or house-keeping genes. Upon binding to their target DNA sequences the forward and reverse probes are ligated. As only bound probes can be ligated, the number of ligated products accurately reflects the abundance of a specific target in the sample being studied. PCR is performed to amplify each product. Each probe pair is designed to be a specific length so that the PCR product relating to each individual target can be separated by capillary electrophoresis. Probes can be identified on a capillary sequencer because the forward primer used for PCR is fluorescently labelled. The amount of probe product detected for each target is relative to the amount of target sequence present in the sample. Signals relating to the target regions are then normalised against the signals generated from the reference probes. Adjusted probe signals are then compared against 3 control DNA samples and deletions or duplications of whole exons and/ or genes can be detected.

### **3.5.3 No whole exon or gene deletions or duplications were detected in the gamma-secretase genes *NCSTN*, *PSEN1* and *PSENEN* in this HS cohort**

At least one affected individual from each of the five pedigrees in which no mutations were detected was assessed for large scale deletions and duplications in *NCSTN*, *PSENEN* and *PSEN1* by MLPA (individuals marked with a green star in Figure 3.1, Figure 3.2, Figure 3.3). No whole exon or gene deletions or duplications were detected in these individuals. A representative plot of peak ratios can be seen in Figure 3.6.



**Figure 3.6.** A plot of the normalised MLPA peak ratios in one individual with HS (probes designed to assess copy number of the exons of *PSEN1* and *PSENEN*). The green squares represent probes designed to bind to exons of interest (relating to *PSEN1* and *PSENEN*). The blue squares represent control probes. The green lines represent the upper and lower thresholds above and below which results would be interpreted as a duplication or deletion respectively. All probes relating to *PSEN1* and *PSENEN* and all control probes fall within the normal limits, thus no whole exon or whole gene duplications or deletions were detected.

### 3.6 Discussion

In light of data from China implicating the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1* in HS pathogenesis we sought to characterise the involvement of mutations in those genes in 7 multigenerational British kindreds demonstrating autosomal dominant inheritance. PCR and direct sequencing of those genes in that cohort revealed that affected members of two of the seven pedigrees had novel mutations in *PSENEN* (an insertion, c.66\_67insG) and *NCSTN* (a splice site substitution, c.1125+1 G>A). None of the remaining five pedigrees harboured rare (<1% population frequency) variants in those genes. It was therefore hypothesised that whole exon/ gene deletions or duplications may underlie HS in the remaining five cases. MLPA was performed in at least one affected individual from each of these pedigrees to interrogate this hypothesis. No CNV was detected.

This was the first report of gamma-secretase mutations in Caucasian individuals and these results provided further evidence that mutations in the gamma-secretase genes *NCSTN* and *PSENEN* underlie some familial cases of HS. It was notable however that no rare variation was detected in those genes in the remaining five kindreds. These findings were corroborated by a more recent study in which only 3 of 14 French pedigrees with HS were found to harbour gamma-secretase mutations in *NCSTN*, *PSEN1* or *PSENEN*<sup>139</sup>. The pedigrees in which no mutations were identified may represent alleles comprising undetected coding variants or non-coding variants, phenocopies of HS or be explained by further locus heterogeneity in HS. One of the limitations of this genetic study was that it only assessed coding regions and associated splice site boundaries, thus variants in non-coding regions such as promoter sequences, enhancer sequences and silencer sequences would not have been detected.

The functional characterisation of the two mutations identified in these British pedigrees became the focus of further study and this is discussed in Chapter 4.

## **Chapter 4: Results: Functional characterisation of the gamma-secretase mutations identified in *PSENEN* and *NCSTN* in familial HS**

### **4.1 Introduction**

Familial genetic studies confirmed the presence of heterozygous gamma-secretase gene mutations in two Caucasian multiplex kindreds with HS (*NCSTN* c.1125+1 G>A in pedigree 3 and *PSENEN* c.66\_67insG in pedigree 4, Figure 3.2).

#### **4.1.1 Aims and objectives**

The aim of this study was to functionally characterise those two mutations. Experiments were performed to identify the effect of each mutation on their respective transcript, protein and on overall gamma-secretase enzyme activity. The location of NCSTN and PEN-2 in the skin of individuals harbouring the mutations and healthy volunteers was examined via immunohistochemistry.

## 4.2 Patient cohort and samples

Samples were obtained from individuals III:1 in pedigree 3, III:2 in pedigree 4 (Figure 3.2) and 3 age-matched healthy volunteer subjects (demographic and clinical details shown in Table 4.1). Each individual provided two 6mm skin biopsies, one from axillary skin (affected region) and one from para-axillary skin (unaffected region). Primary human fibroblasts were grown in culture and used to harvest RNA, whole cell protein and solubilised membrane protein. Whole skin sections were examined by immunohistochemistry.

**Table 4.1. Demographics and clinical details of the two mutation-positive patients and three healthy volunteer subjects.**

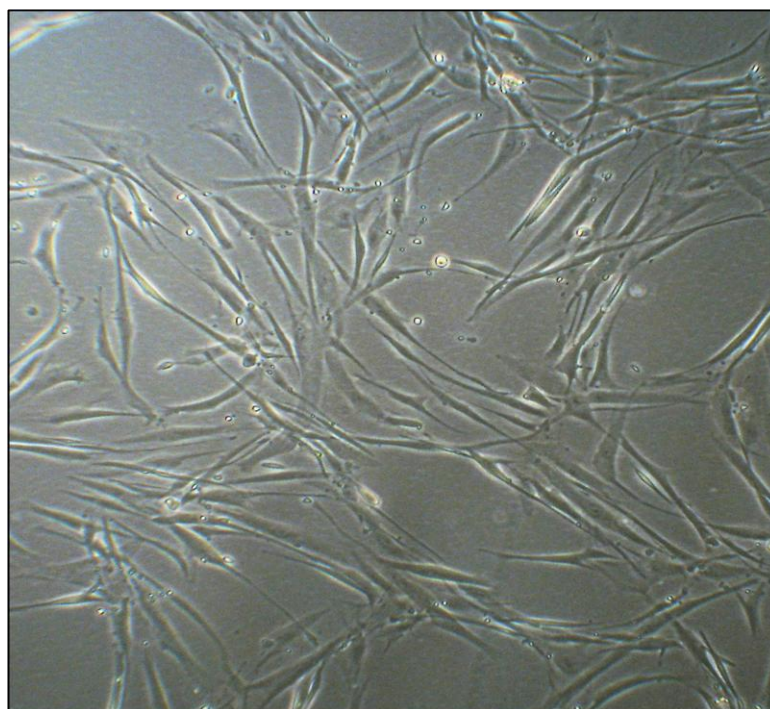
	<b>Individual with <i>NCSTN</i> mutation</b>	<b>Individual with <i>PSENEN</i> mutation</b>	<b>Healthy volunteers (3)</b>
<b>Age</b>	46	41	37, 39, 35 (average 37)
<b>Sex</b>	M	F	M, F, F
<b>Ethnicity</b>	White Caucasian	White Caucasian	White Caucasian
<b>Smoking status</b>	Ex-smoker	Non-smoker	Ex-smoker, 2 non-smokers
<b>BMI</b>	30	23	22, 25, 24 (average 23.8)

### **4.3 *NCSTN* transcript and NCSTN protein analysis in an individual harbouring the heterozygous *NCSTN* splice site substitution *NCSTN* c.1125+1 G>A**

#### **4.3.1 *NCSTN* c. 1125+1 G>A results in the skipping of *NCSTN* exon 9**

##### **4.3.1.1 PCR and agarose gel electrophoresis of the full length *NCSTN* transcript**

RNA was extracted from primary human fibroblasts (Figure 4.1) and reverse transcribed (RT-PCR) to produce cDNA. cDNA is more stable than RNA and thus facilitates experimental analysis of gene transcripts.



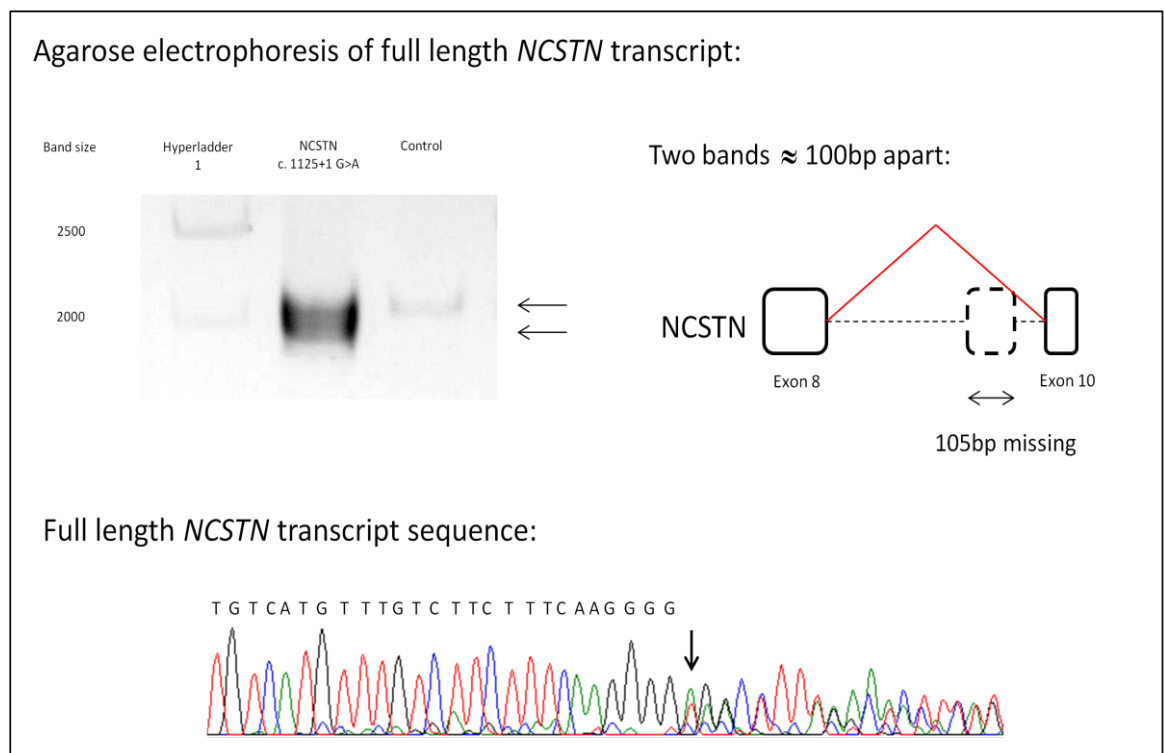
**Figure 4.1. Primary human fibroblasts grown in culture.**

Primers were designed to amplify the full length *NCSTN* transcript (full length 2.9Kb, amplified segment 2Kb). PCR amplification required optimisation. The efficacy of each reaction was assessed via agarose gel electrophoresis of the PCR product. cDNA from the mutation-positive patient and a healthy volunteer was amplified. One band was

observed at the expected size (2Kb) in the healthy volunteer (control). Two bands were observed in the mutation-positive patient, the expected band at 2Kb and an additional band at 1.9Kb, suggesting the mutation leads to the generation of a truncated mutant transcript (Figure 4.2).

#### **4.3.1.2 Direct sequencing of the full length NCSTN transcript**

Full length NCSTN cDNA from the mutation-positive patient and the healthy volunteer was sequenced to determine the nature of the mutant transcript. Five pairs of overlapping primers were designed to sequence the 2Kb transcript. Sequencing revealed that *NCSTN* exon 9 is missing in the mutant allele (Figure 4.2). *NCSTN* exon 9 is 105 bases long which approximately corresponds with the distance observed between the wild type (WT) and mutant band on agarose gel electrophoresis (Figure 4.2).



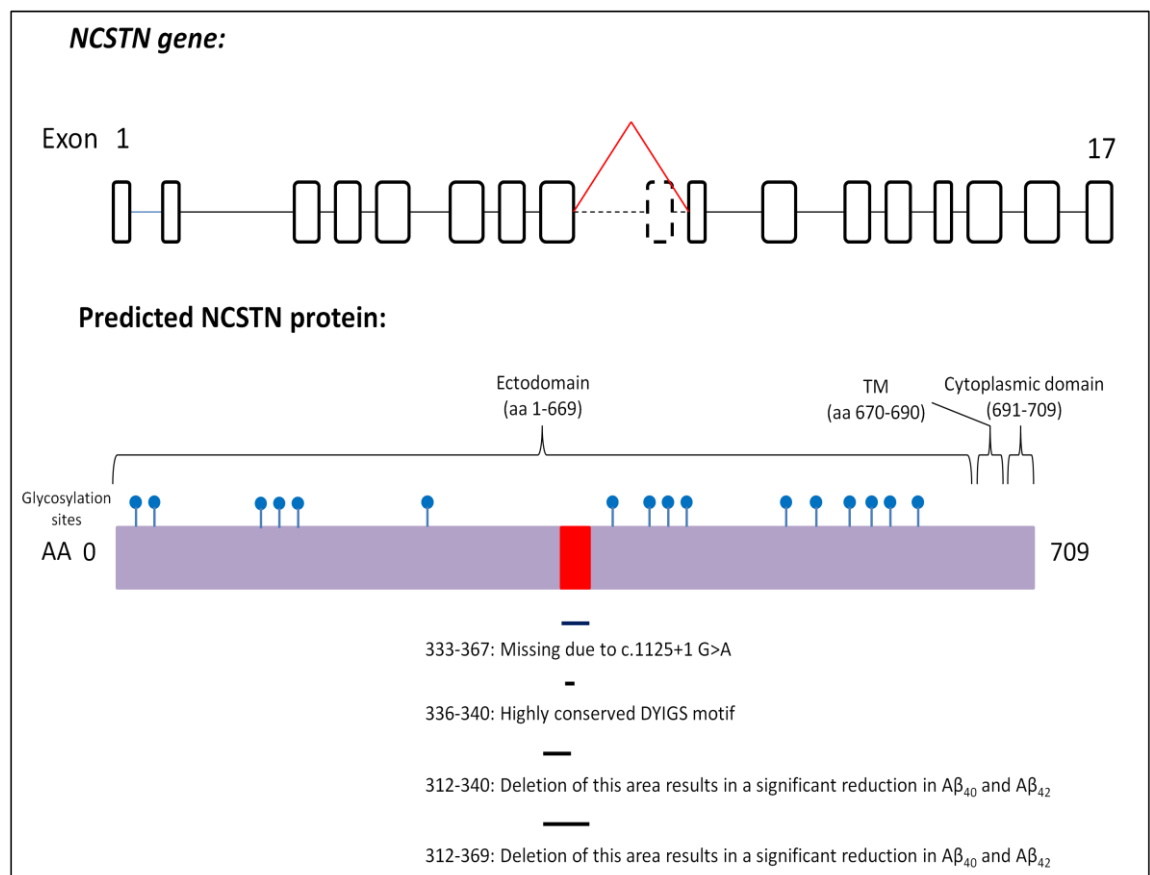
**Figure 4.2. Agarose electrophoresis and Sanger sequencing of the full length *NCSTN* transcript in the patient harbouring the heterozygous *NCSTN* 1125+1 G>A mutation.** On electrophoresis, two bands were seen in the patient harbouring the mutation versus one band in a healthy volunteer. The additional band migrated further and appeared to represent a transcript that was around 100 bases lighter. Sequencing corroborated that finding, demonstrating skipping of *NCSTN* exon 9 in the mutant transcript.

#### 4.3.1.3 The *NCSTN* c.1125+1 G>A mutation is predicted to result in a truncated *NCSTN* protein

*NCSTN* exon 9 is 105 bases long (divisible by three), thus any absence of this sequence in the mutant transcript would not alter the reading frame but would theoretically result in a truncated protein product (amino acids 333-367 predicted to be missing, Figure 4.3). Exon 9 encodes part of the large nicastrin ectodomain which is known to be integral to normal gamma-secretase function. Three domains within the Nicastrin protein are highly conserved between animal species, amino acids 306 – 360, 419 – 458 and 625 – 662. The amino acids affected by this mutation fall within the first of these regions within which there is a very highly conserved DYIGS motif (aa 336 – 340)<sup>171</sup>.



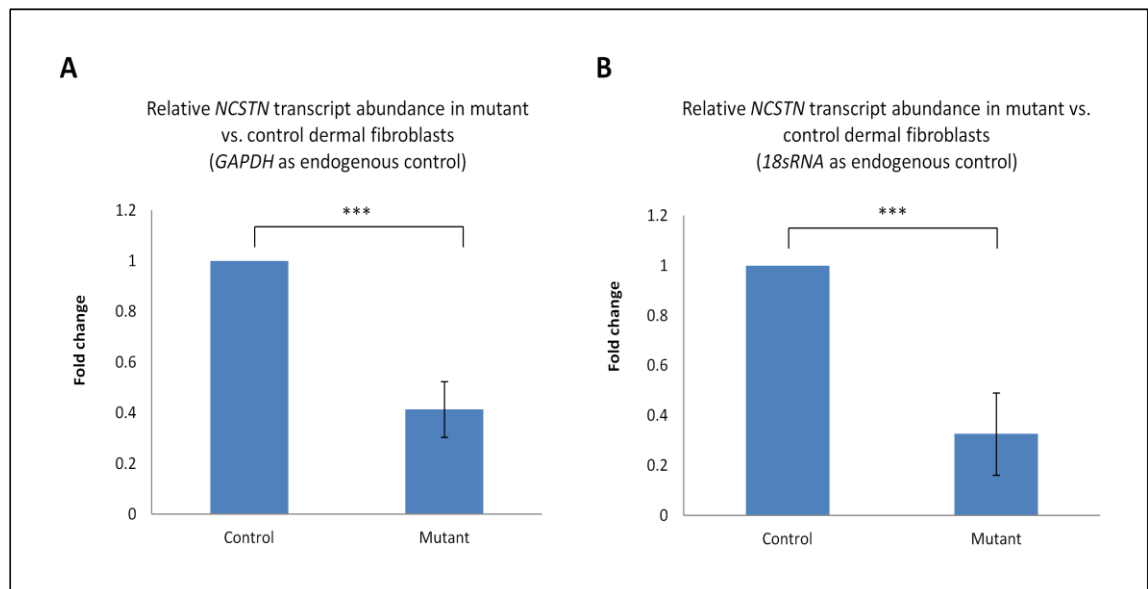
Missense mutation of this motif (to AAIGS) has been shown to result in an increase in A $\beta$  secretion following APP cleavage whereas homozygous deletion of residues 312-340 and 312-369 results in a significant reduction in A $\beta_{40}$  and A $\beta_{42}$  production. This effect was more pronounced with the residue 312-369 deletion, which is most similar to the protein predicted to arise as a consequence of this splice site mutation<sup>171</sup>. This data strongly supports a pathological role for this mutation and suggests that, if the mutant transcript is translated, the resultant mutant protein could significantly affect gamma-secretase function.



**Figure 4.3. The consequences of the *NCSTN* c.1125+1 G>A mutation on the *NCSTN* transcript and protein.** This mutation results in defective splicing and skipping of *NCSTN* exon 9. This in turn is predicted to result in a truncated NCSTN ectodomain. The region of the ectodomain affected is highly conserved and mutagenesis experiments affecting the highlighted amino acid sequences significantly affect gamma-secretase carboxypeptidase activity, reducing the overall amount of A $\beta_{40}$  and A $\beta_{42}$  produced following APP cleavage.

#### **4.3.2 *NCSTN* transcript abundance is reduced as a consequence of *NCSTN* c.1125+1 G>A**

Transcript abundance in the mutation-positive individual and three healthy volunteers (controls) was assessed by real time quantitative PCR. RNA was harvested from primary human fibroblasts (grown from skin taken from the affected area in the patient harbouring the mutation and matched axillary skin in volunteer subjects) and equal concentrations were reverse transcribed to produce cDNA. Real time quantitative PCR was performed using Taqman gene expression assays (Life Technologies). The principle of these assays is that a fluorescent taqman probe binds to a specific part of the transcript. Each time the transcript replicates during the real time PCR fluorescence is released and measured. The amount of fluorescence emitted therefore reflects the quantity of transcript present in the initial reaction. This result is controlled against a gene transcript with reliable and stable expression, otherwise termed a housekeeping gene, as it cannot be assumed that equal amounts of cDNA are loaded in every reaction. A commercially available *NCSTN* Taqman probe was chosen for this experiment that binds to an exon 2 related portion of the *NCSTN* transcript. This is well before the site of the *NCSTN* c. 1125+1 G>A mutation, thus the assay results reflect total *NCSTN* transcript abundance (WT and mutant combined) rather than that of the wild type or mutant transcript alone. Transcript expression was significantly reduced in mutant fibroblasts versus those harvested from healthy controls (Figure 4.4). Experiments were initially run with *GAPDH* as an endogenous control (Figure 4.4a). To confirm the validity of the results generated, repeat assays were run with a second endogenous control, *18sRNA* (Figure 4.4b). Results were consistent with those generated using *GAPDH*. Consequently, *GAPDH* was used as the endogenous control for all further real time PCR reactions reported in this thesis.



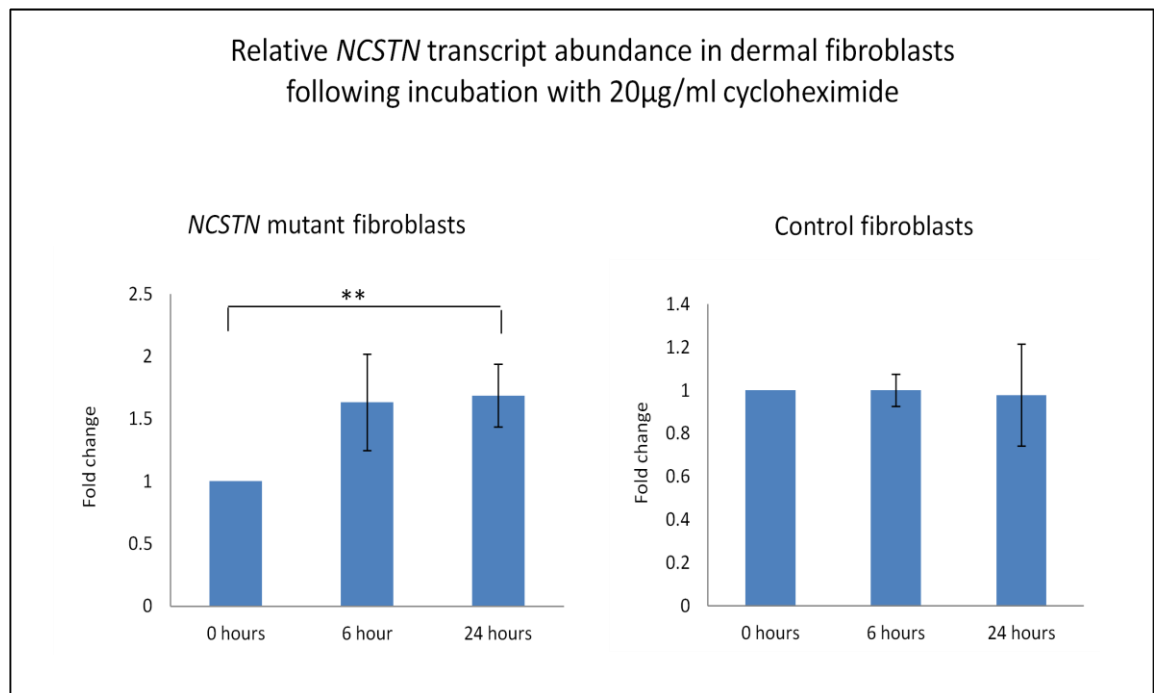
**Figure 4.4. Relative *NCSTN* transcript abundance in the individual harbouring *NCSTN* c.1125+1 G>A versus healthy volunteers (controls).** All real time PCR reactions were performed in triplicate and experiments were repeated three times. The fold changes represent the average of the three experiments A) *GAPDH* was used as an endogenous control. There is a significant reduction in transcript abundance in the mutation-positive patient versus control. The figure plots mutant expression against one control subject however the result was representative of experiments using two further independent controls ( $p < 0.01$ ,  $p < 0.01$ ). B) *18sRNA* used as an endogenous control. Again, there was a significant reduction in transcript abundance in the mutation-positive patient versus control and this result was representative of results obtained with two further controls ( $p < 0.001$ ,  $p < 0.001$ ) (error bars = standard deviation (SD), \*\*\* denotes  $p < 0.001$ , unpaired t-test).

### 4.3.3 The mutant *NCSTN* transcript is subject to decay

Real time PCR data suggested that overall *NCSTN* transcript abundance was reduced in mutant fibroblasts versus controls. It was therefore hypothesised that the mutant transcript is subject to decay. Numerous mRNA surveillance mechanisms have evolved in humans to detect and eliminate aberrant transcripts. These mechanisms are active in both the nuclear and extra-nuclear compartment. The nuclear exosome screens pre-mRNA for splicing defects and appropriate processing prior to release from the nucleus<sup>269</sup>. Within the extra-nuclear compartment, mRNA surveillance appears to be most active at the stage of translation. Three specific surveillance mechanisms have

been reported, nonsense-mediated decay (NMD, this mechanism recognises premature stop codons and induces transcript decay), no-go decay (this mechanism recognises when ribosomes get stuck processing transcripts and induces transcript decay) and non-stop decay (a mechanism that is activated when no stop codon is detected)<sup>270</sup>.

Methods of investigating nuclear mechanisms of RNA decay (nuclear exosome) are limited however extra-nuclear decay can be interrogated by exposing cells to cycloheximide. Cycloheximide is a chemical that can block protein translation by inhibiting translational elongation. Consequently it has been shown to inhibit the translation related RNA surveillance mechanism NMD<sup>267</sup>. Mutant and control fibroblasts were exposed to 20µg/ml cycloheximide for up to 24 hours. An incremental increase in transcript abundance was observed in cycloheximide treated cells over the 24 hour period (Figure 4.5). This would suggest that the mutant transcript is subject to decay and that the mechanism of RNA decay is arising at the level of translation. As previously discussed, the splice site mutation appears to result in the skipping of NCSTN exon 9 but is not predicted to alter the reading frame or cause premature truncation, thus NMD is an unlikely mechanism of RNA decay. The mutation is not predicted to disrupt the native stop codon and would not therefore seem likely to induce no-stop decay. It is interesting however that upon running the WT and mutant RNA sequences through an RNA secondary structure prediction website (<http://rna.urmc.rochester.edu/RNAstructureWeb/>) that the mutant RNA is predicted to adopt an altered secondary structure (Figure 4.6). This may in itself confer instability and result in subsequent decay but the association of that hypothetical mechanism with translation is unknown. A further explanation is that the altered RNA structure may affect ribosomal processing and potentially induce no-go decay<sup>271</sup>. The exact mechanism of RNA decay remains to be elicited.

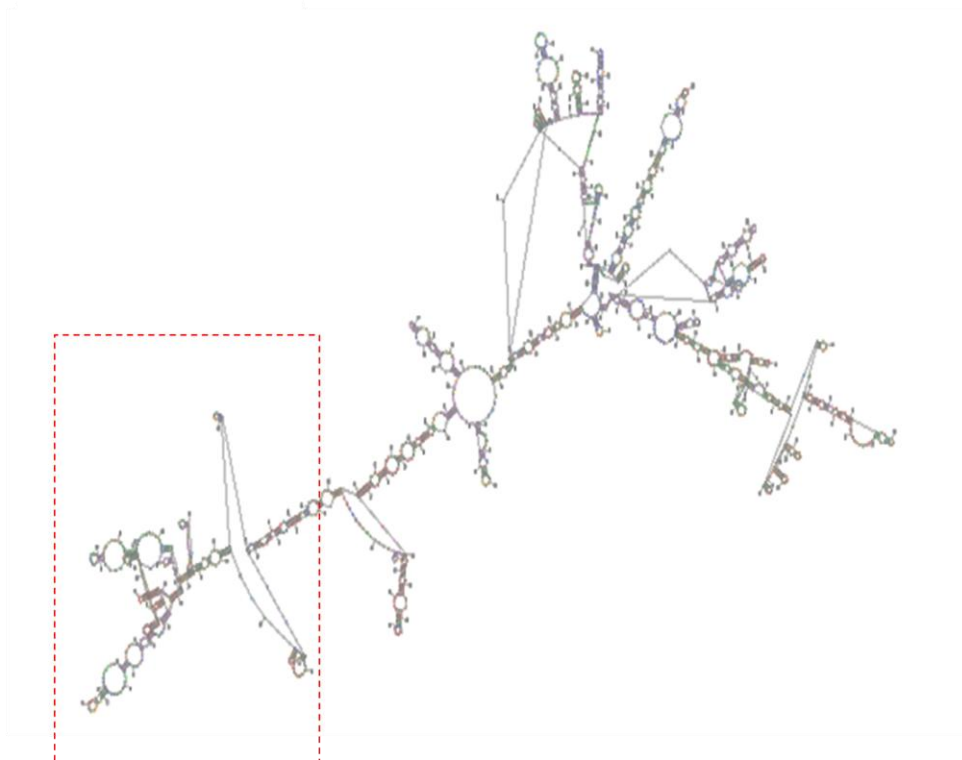


**Figure 4.5. *NCSTN* transcript abundance within mutant and control fibroblasts following incubation with cycloheximide.** There was a significantly increase in *NCSTN* transcript abundance following a 24 hour incubation with cycloheximide. This was not the case with control fibroblasts. All experiments were performed in triplicate and repeated three times. Fold changes represent the average fold change achieved over all three experiments (error bars = SD, \*\* denotes  $p < 0.01$ , unpaired t-test).

WT *NCSTN* transcript



Mutant *NCSTN* c.1125+1 G>A transcript



**Figure 4.6. Predicted RNA secondary structure of the WT and mutant (c.1125+1 G>A) *NCSTN* transcript.** These images represent the lowest free energy structure for each respective RNA sequence. The *NCSTN* c.1125+1 G>A mutation results in a truncated RNA transcript,

missing the section corresponding to *NCSTN* exon 9. This is predicted to alter the folding and orientation of the secondary RNA structure (lower left portion as represented here, highlighted by red box). The images were taken directly from an RNA secondary structure prediction website (<http://rna.urmc.rochester.edu/RNAstructureWeb/>).

#### **4.3.4 Nicastrin protein expression is reduced in mutant fibroblasts versus control fibroblasts**

*NCSTN* transcript expression was reduced in the individual carrying the *NCSTN* c.1125+1 G>A mutation, thus it was hypothesised that there would be a corresponding reduction in protein expression. Given that some mutant transcript was present (as seen on agarose gel electrophoresis and upon sequencing the full length *NCSTN* transcript), a further hypothesis was that there would be some expression of a mutant form of nicastrin. Total cell protein was extracted from primary human fibroblasts in culture (chemical lysis with RIPA buffer) and immunoblotting was performed to interrogate these hypotheses.

##### **4.3.4.1 Antibody selection and optimisation**

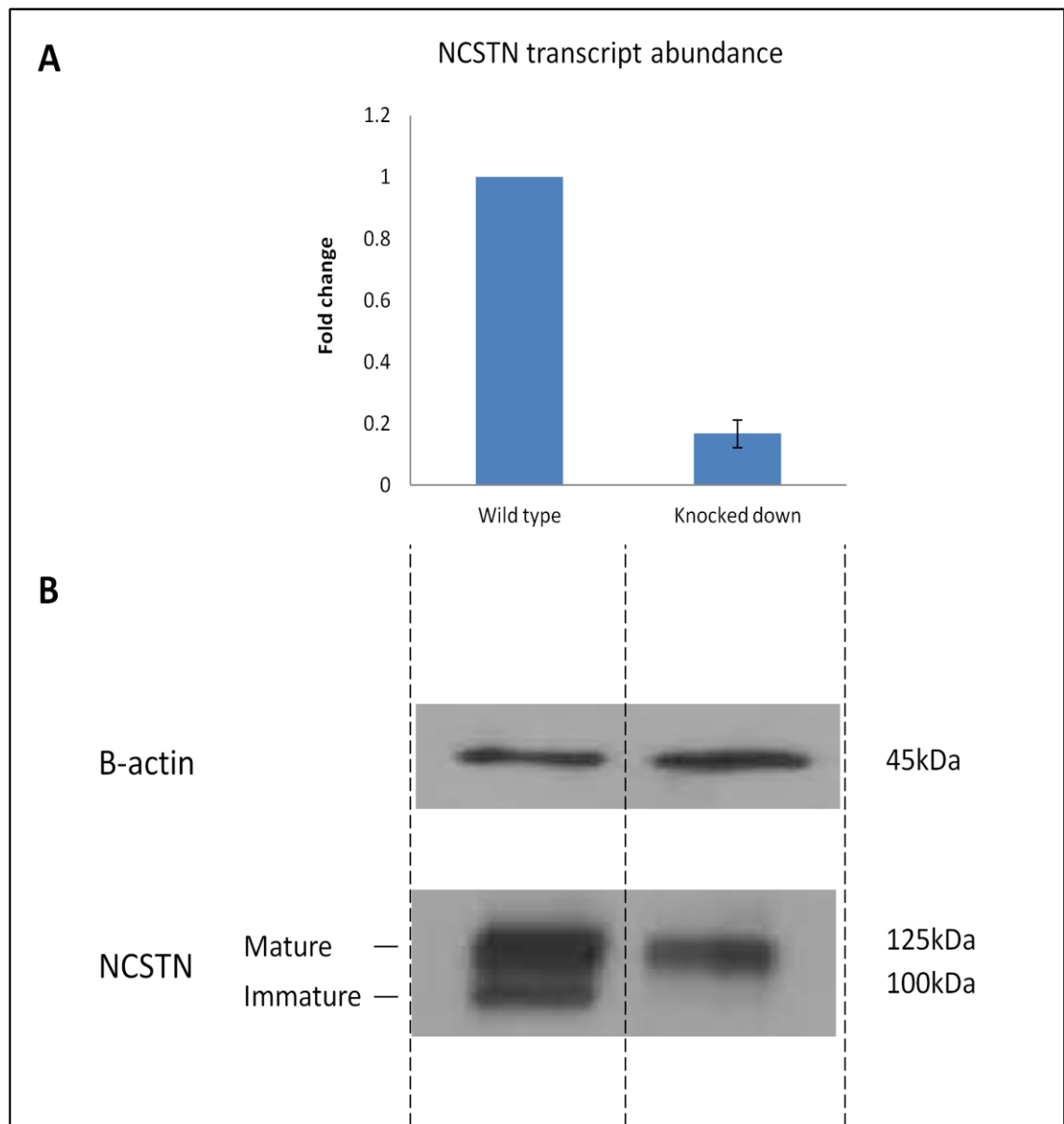
Four rabbit polyclonal anti-nicastrin antibodies were tested for specificity and sensitivity. An anti-nicastrin antibody was chosen that produced few non-specific bands, had successfully been used to perform immunoblotting and immunohistochemistry in previous publications and because it binds to the N-terminus of Nicastrin (immunogen CFTRDLMEKLKGRTSRIAGLAVSL, aa 103 – 124). The binding site is well before the segment of nicastrin related to the mutated portion of *NCSTN* (relating to aa 332 – 368), thus the antibody would theoretically detect all nicastrin protein variants present (wild type and mutant protein). Conversely, an antibody that binds at or after amino acids 332 may only detect the wild type protein. The total amount of protein loaded for gel electrophoresis was optimised, 20µg was sufficient to produce strong and specific bands with an antibody titre of 1:400. As

discussed in the introduction, nicastrin exists in a mature and immature form in cells. It therefore generates two bands on immunoblotting, a light immature band and a heavier mature band<sup>198</sup>. Protein electrophoresis on a 10% Tris-HCL gel sufficiently separated those two bands.



#### **4.3.4.2 SiRNA knockdown of *NCSTN* in primary human fibroblasts confirmed the specificity of the nicastrin bands observed on immunoblotting**

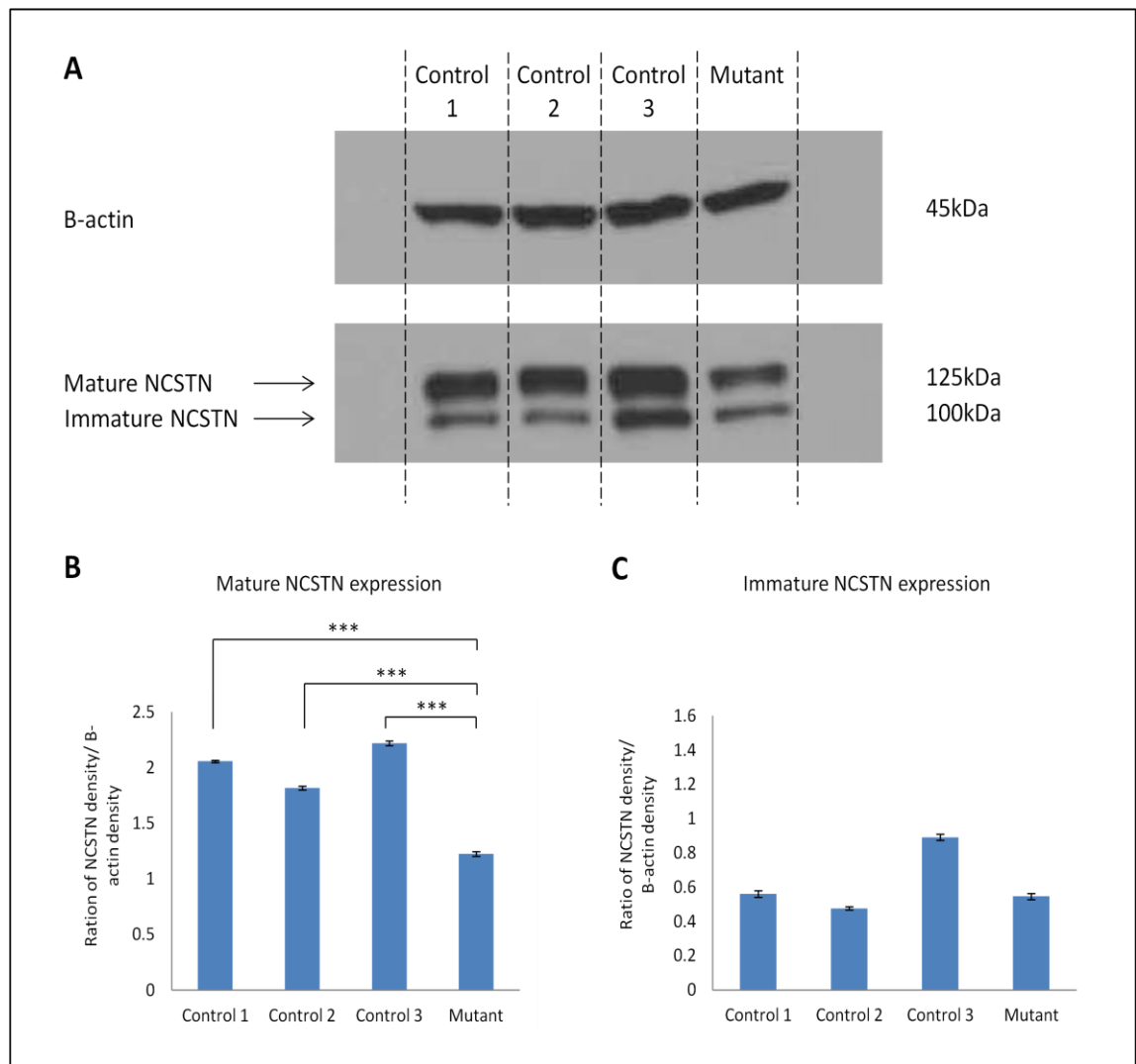
The predicted molecular weight of the nicastrin band, as detailed by the manufacturers of the chosen antibody, was 85kDa however the immature and mature bands were running at 100kDa and 125kDa on optimisation blots. Despite differing from the manufacturers prediction, the appearance and apparent molecular weight of the two protein bands was entirely consistent with previous publications<sup>272; 273</sup>. Given the discrepancy, SiRNA knockdown experiments were performed to confirm the specificity of the bands observed. Primary human fibroblasts derived from a healthy volunteer were transfected with commercially available *NCSTN* SiRNA and cultured for 48 hours. Untreated fibroblasts from the same patient were cultured in parallel. RNA and protein were extracted from both the transfected and un-transfected cells. RT-PCR of equal quantities of RNA produced cDNA. Real-time quantitative PCR of the cDNA products confirmed that sufficient knockdown of *NCSTN* transcript expression had occurred in the SiRNA treated cells (Figure 4.7). The total protein concentration harvested from treated and un-treated cells was calculated using a BCA assay. 20µg of protein from treated and untreated cells was immunoblotted and there was a clear reduction in the blot intensity of the nicastrin bands (mature and immature) in SiRNA transfected fibroblasts versus un-transfected fibroblasts (Figure 4.7).



**Figure 4.7. NCSTN transcript abundance and NCSTN protein expression in primary human fibroblasts following SiRNA knockdown.** A) The efficacy of knockdown was shown by quantitative RT-PCR. There was a significant reduction in transcript abundance in the SiRNA treated cells compared to control cells. All quantitative RT-PCR experiments were performed in triplicate and repeated three times. The fold change represents the average fold change observed over the three experiments (error bar = SD, \*\*\* denotes  $p < 0.001$ , unpaired t-test). B) An immunoblot showing a corresponding reduction in NCSTN protein (mature and immature) expression in SiRNA treated cells. Immature protein expression was affected to a greater extent given the short 48 hour incubation with SiRNA.

#### **4.3.4.3 Nicastrin expression was reduced in mutant fibroblasts versus control fibroblasts**

Total cell protein was harvested from primary human fibroblasts from the mutation-positive patient and three healthy volunteers. The total protein concentration in each sample was measured by BCA assay. 20µg total cell protein was immunoblotted for each sample and β-actin was used as a loading control. Differences in protein expression were quantified from each blot using Image J (National Institutes of Health (NIH), USA) which assessed the blot density of each sample. All nicastrin band densities were controlled against their corresponding β-actin blot density to control for loading variation. There was a significant reduction in mature nicastrin expression in fibroblasts harvested from the mutation-positive patient versus healthy volunteers ( $p < 0.001$ , Figure 4.8). There was no associated reduction in immature nicastrin expression. No additional protein bands were present in the individual harbouring the mutation to support the presence of a mutant protein. Increasing blot time and the Tris-HCL gel percentage increased the distance between mature and immature bands but did not identify a mutant protein.

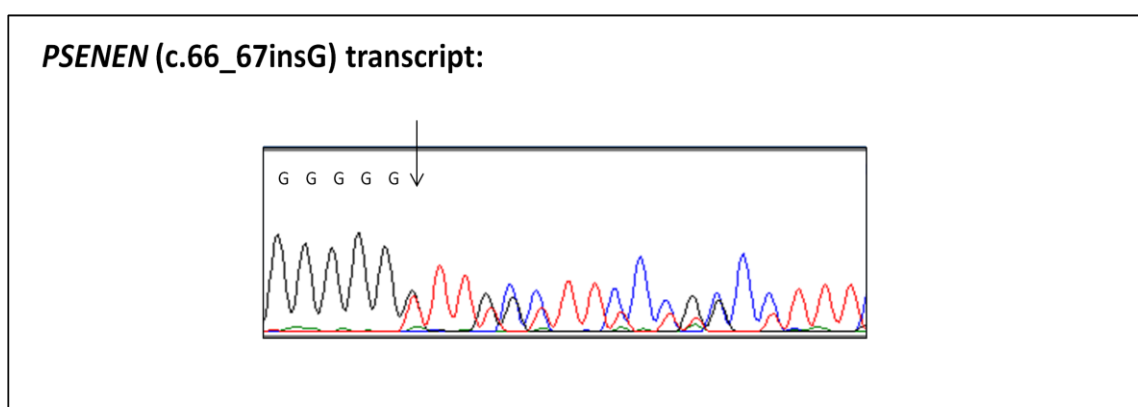


**Figure 4.8. Immunoblot of NCSTN in mutant and control primary human fibroblasts.** A) Immunoblot of NCSTN in mutant and wild type fibroblasts. There appeared to be a significant reduction in mature NCSTN expression in the mutation-positive patient compared to healthy volunteers (controls).  $\beta$ -actin was blotted as an endogenous control and loading appeared equal. B, C) Protein band densities were measured and controlled against the blot density of the corresponding  $\beta$ -actin band. NCSTN/ B-actin band densities were plotted as a measure of total NCSTN expression in each sample. There was a significant reduction in mature protein expression in mutant fibroblasts versus control fibroblasts but no significant difference in immature protein expression (error bars = SD, \*\*\* denotes  $p < 0.001$ , unpaired t-test).

#### **4.4 *PSENEN* transcript and *PSENEN* protein analysis in an individual harbouring the heterozygous *PSENEN* insertion *PSENEN* c.66\_67insG**

##### **4.4.1 *PSENEN* c.66\_67insG results in a frameshift and is predicted to result in an altered and lengthened protein product**

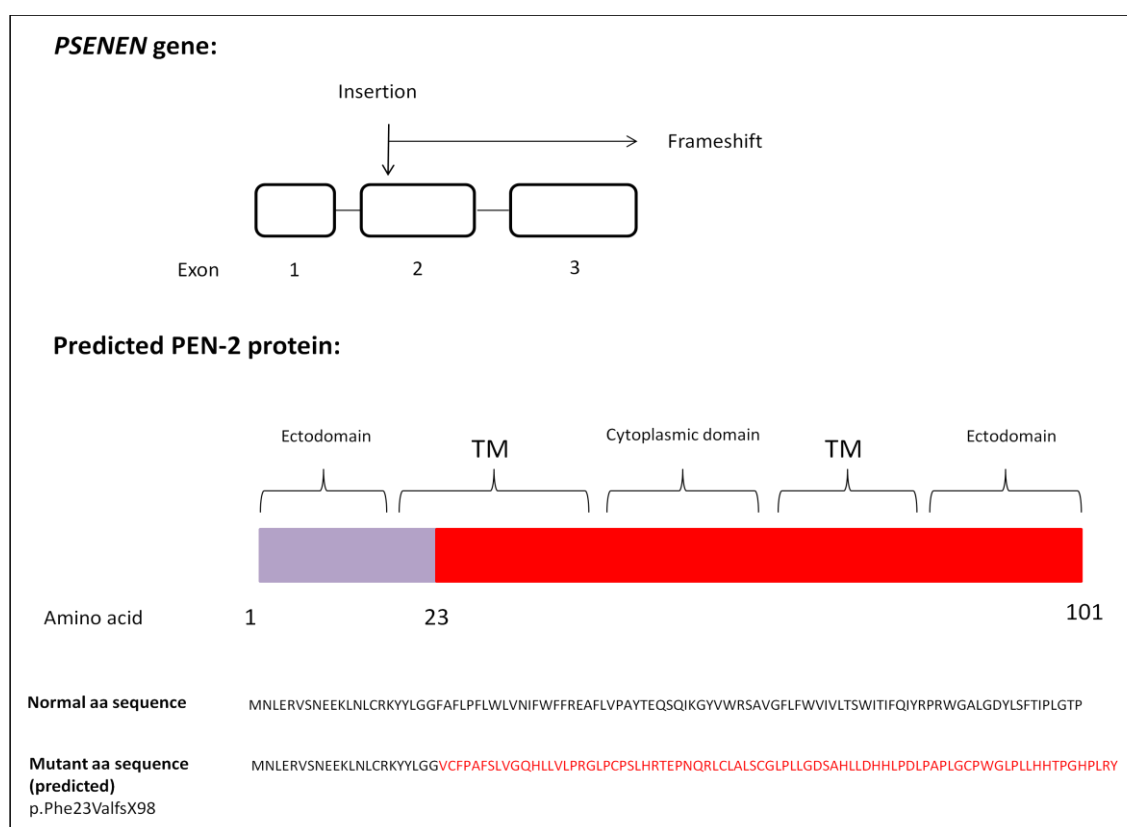
Primary human fibroblasts were cultured from an individual with the *PSENEN* c.66\_67insG mutation. RNA was harvested from these cells and converted to cDNA by RT-PCR. PCR primers were designed to bind to regions of cDNA flanking the corresponding exonic location of this mutation (mutation within exon 2, primers designed to bind to sequences corresponding to exons 1 and 3). Amplification and direct sequencing of this segment demonstrated that the insertion is present in the transcript (Figure 4.9).



**Figure 4.9. *PSENEN* cDNA sequence in an individual harbouring the heterozygous c.66\_67insG.** The insertion of a guanine base at position 67 is indicated by the arrow.

This single base insertion is predicted to alter the reading frame and subsequently result in an altered protein product (Figure 4.10). PEN-2 has two ectodomains, two transmembrane domains and a cytoplasmic domain. This mutation is predicted to alter amino acids 23 through to 104 of PEN-2 where there is predicted to be a newly formed

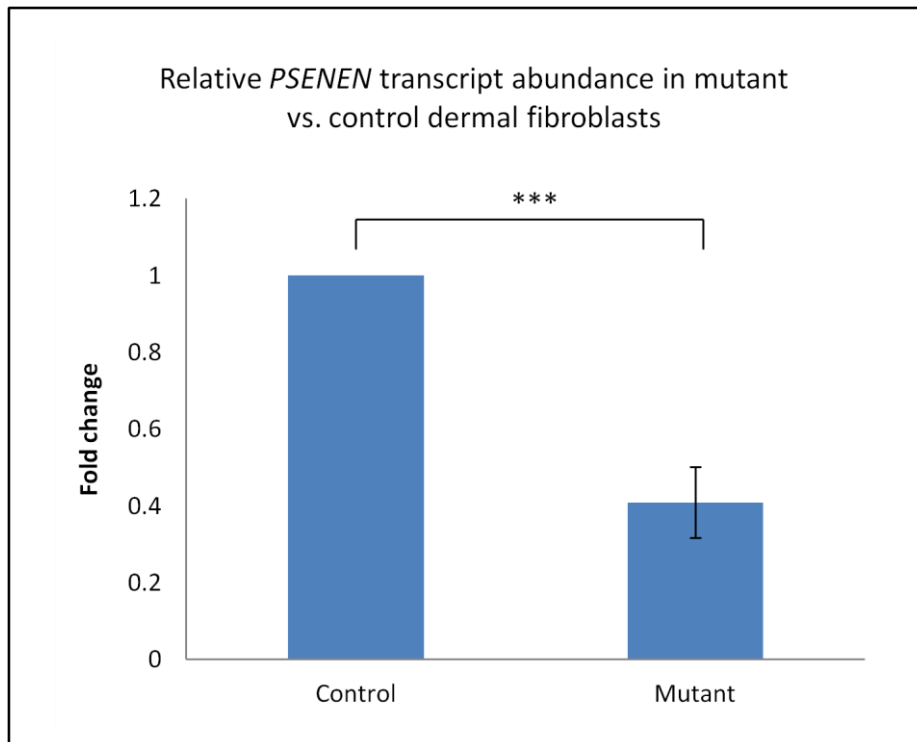
stop codon (p.Phe23ValfsX98). The altered protein is predicted to be three bases longer than wild type PEN-2 (101aa). This mutation would therefore affect the two trans-membrane domains, the cytoplasmic domain and one ectodomain, leading one to predict that it could have a significant impact on protein function. The proximal two-thirds of the first trans-membrane domain of PEN-2 is known to be critically important for facilitating the endoproteolysis of presenilin-1<sup>208</sup>. Furthermore, the carboxy-terminal end of PEN-2 is critical in stabilising the presenilin 1 derivatives created by endoproteolysis<sup>208</sup>.



**Figure 4.10. The predicted *PSENEN* c.66\_67insG mutant transcript and protein.** The inserted guanine lies within exon 2 of *PSENEN* and alters the reading frame. The protein product is therefore predicted to be altered from amino acid 23 onwards and 3 amino acids longer than the wild-type PEN-2 protein.

#### **4.4.2 *PSENEN* transcript abundance is reduced as a consequence of *PSENEN* c.66\_67insG**

*PSENEN* transcript abundance in the mutation-positive patient and three healthy volunteers (controls) was assessed by real time quantitative PCR. RNA was harvested from primary human fibroblasts cultured from affected skin in the mutation-positive patient and a matched axillary site in the three volunteer subjects. cDNA was produced from equal concentrations of RNA via RT-PCR. Real time quantitative PCR was performed using Taqman gene expression assays (Life Technologies). A specific commercially available *PSENEN* Taqman probe was chosen that binds to an exon 1 related portion of the *PSENEN* transcript. This is before the site of the c.66\_67insG mutation, thus the assay results reflect total *PSENEN* transcript abundance (WT and mutant combined) rather than that of the wild type or mutant transcript alone. Transcript expression was significantly reduced in mutant versus control fibroblasts ( $p < 0.001$ , Figure 4.11).



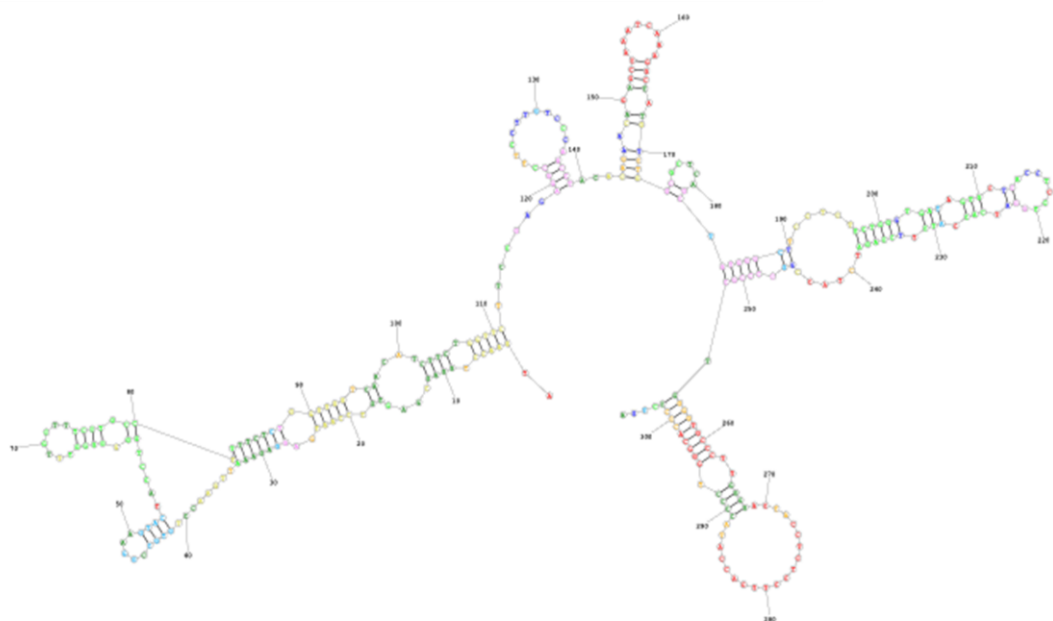
**Figure 4.11. Relative *PSENEN* transcript abundance in mutant and control primary human fibroblasts.** *PSENEN* transcript abundance was significantly reduced in mutant fibroblasts versus control fibroblasts. The figure plots transcript expression in the mutation-positive patient versus one healthy volunteer (control) however this result was representative of experiments using two further independent controls ( $p < 0.05$ ,  $p < 0.05$ ). All real time PCR reactions were performed in triplicate and experiments were repeated three times. The fold changes represent the average of the three experiments (error bars = SD, \*\*\* denotes  $p < 0.001$ , unpaired t-test).

#### 4.4.3 The mechanism underlying the apparent reduction in *PSENEN* transcript abundance is unclear

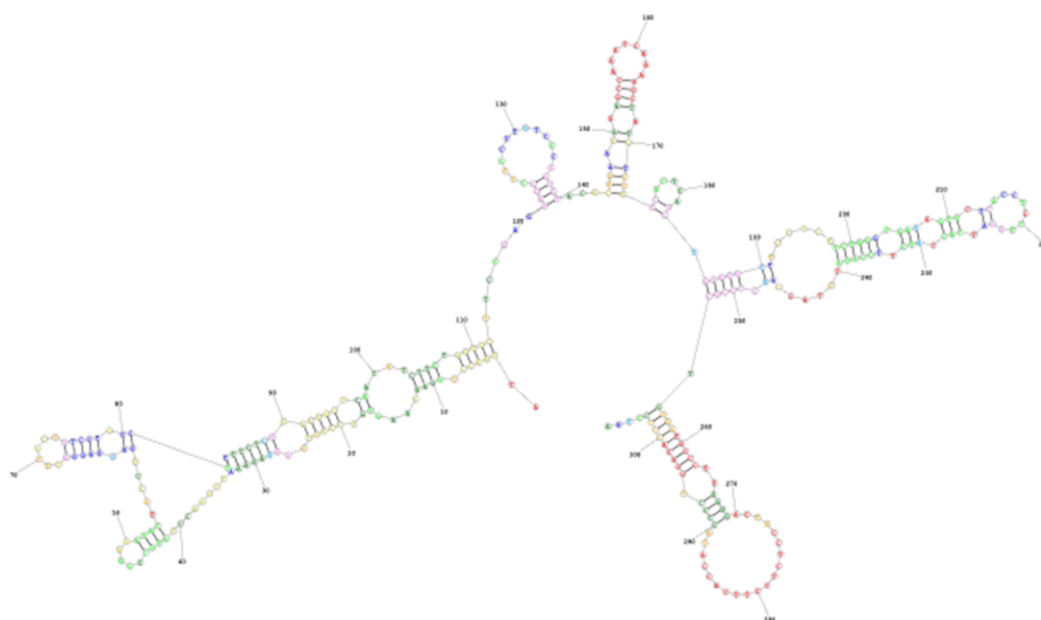
*PSENEN* transcript abundance was reduced in the individual harbouring the mutation versus healthy volunteers suggesting that the mutant transcript may be subject to decay. WT and mutant *PSENEN* transcripts were run through an RNA secondary structure prediction website (<http://rna.urmc.rochester.edu/RNAstructureWeb/>). No significant alteration in secondary structure was evident as a consequence of the single base insertion (Figure 4.12). This would imply that structural instability is an unlikely cause of the apparent reduction in transcript abundance in the patient harbouring the mutation.



### WT *PSENER* transcript



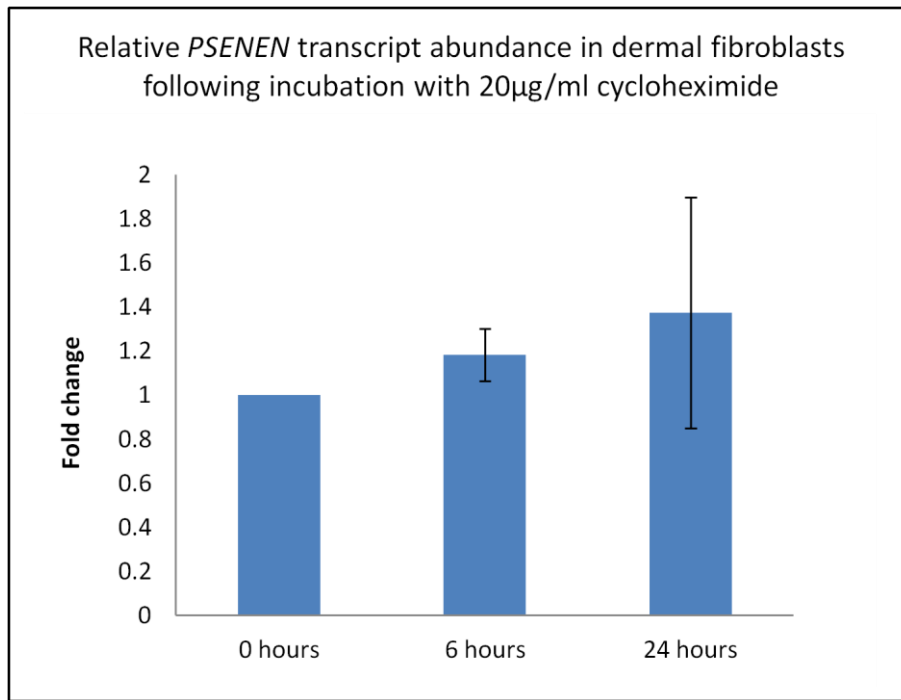
### Mutant *PSENER* c.66\_67insG transcript



**Figure 4.12. Predicted RNA secondary structure of the WT and mutant (c.66\_67insG) *PSENER* transcript.** These images represent the lowest free energy structure for each respective RNA sequence. The secondary structure of the mutated RNA is not predicted to

significantly differ from that of the wild type transcript. Individual images were taken directly from an RNA secondary structure prediction website (<http://rna.urmc.rochester.edu/RNAstructureWeb/>).

To further investigate the mechanism underlying the reduced transcript abundance mutant cells were cultured in 20µg/ml cycloheximide for up to 24 hours. A minimal increase in transcript abundance was observed in the cycloheximide treated cells over the 24 hour period but this was not statistically significant (Figure 4.13). The *PSENEN* mutation was previously shown to result in a frameshift and eliminate the native stop codon. This would theoretically result in the ribosomal machinery reading through the region relating to the native stop codon and may therefore induce non-stop decay (mechanisms of RNA surveillance were discussed in 4.3.3). There would appear to be a newly formed stop codon only ten bases after that site however. Furthermore, non-stop decay is a translation-associated mechanism and as such, one might have expected cycloheximide to increase RNA expression. Whilst this result does not preclude the process of non-stop decay it does infer that further investigation is required to clarify the exact mechanism by which transcript abundance is reduced as a consequence of the *PSENEN* c.66\_67insG mutation.



**Figure 4.13. *PSENEN* transcript abundance in mutant fibroblasts following incubation with cycloheximide.** There was no significant increase in transcript abundance following a 24 hour incubation with cycloheximide. All experiments were performed in triplicate and repeated three times. Fold changes represent the average fold change achieved over all three experiments (error bars = SD).

#### **4.4.4 PEN-2 protein expression was reduced in mutant fibroblasts compared to controls**

*PSENEN* transcript expression was reduced in the individual harbouring the *PSENEN* c.66\_67insG mutation. This led to the hypothesis that there would be a corresponding reduction in protein expression. Given that some mutant transcript is present (seen when sequencing the full length *PSENEN* transcript) a further hypothesis was that there would be some expression of a mutant form of PEN-2. Total cell protein was extracted from primary human fibroblasts in culture (chemical lysis with RIPA buffer) and immunoblotting was performed to interrogate these hypotheses.

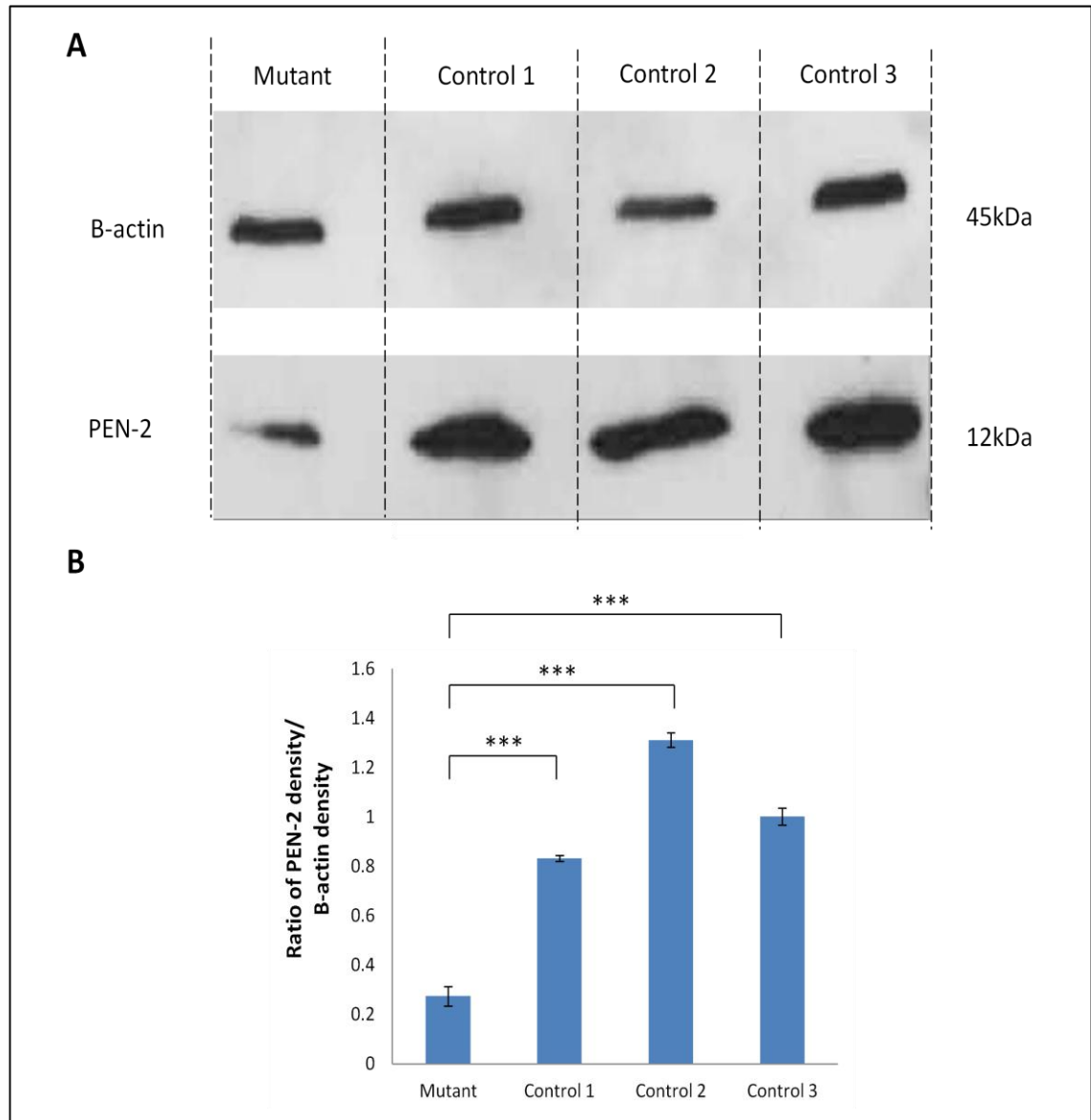
#### **4.4.4.1 Antibody selection and optimisation**

Two rabbit polyclonal anti-PEN-2 antibodies were tested for specificity and sensitivity. An antibody was chosen on the basis that it produced a specific band corresponding to the expected protein weight (12kDa, consistent with previous publications and the use of different PEN-2 antibodies<sup>274</sup>) and because it binds to the N-terminus of PEN-2. The exact immunogen was commercially sensitive however the binding site was confirmed by the manufacturer to be before the segment of PEN-2 corresponding to the insertion (relating to aa 23). The antibody would therefore detect all PEN-2 protein variants present (wild type and mutant protein). The total amount of protein required for each blot was optimised. 25µg was sufficient to produce strong and specific bands with an antibody titre of 1:200. Protein electrophoresis was performed on a 4-20% Tris-HCL gel.

#### **4.4.4.2 PEN-2 expression was reduced in mutant fibroblasts versus control fibroblasts**

Total cell protein was extracted from primary human fibroblasts cultured from the patient harbouring the mutation and three healthy volunteers (controls). The total protein concentration in each sample was measured by BCA assay. 25µg total cell protein was immunoblotted for each sample and β-actin was used as a loading control. Differences in protein expression were quantified from each blot using Image J (NIH) which assessed the blot density of each sample. Whilst the β-actin bands in all samples appeared very similar, all PEN-2 band densities were controlled against their corresponding β-actin blot density for complete accuracy. There was a significant reduction in PEN-2 expression in dermal fibroblasts harvested from the mutation-positive patient versus those harvested from three control subjects (Figure 4.14). No second band was observed in the patient harbouring the mutation to support the presence of a mutant protein. This should be interpreted with caution however given

that the mutant protein is only predicted to be three amino acids longer in length and would therefore be difficult to differentiate from the wild type protein on immunoblotting.



**Figure 4.14. Immunoblot of PEN-2 in protein samples harvested from mutant and control primary human fibroblasts.** A) Immunoblot of PEN-2 in mutant and control primary human fibroblasts. PEN-2 expression appears reduced in mutant fibroblasts versus controls. B-actin was blotted as a loading control and loading appeared equal B) Plot of PEN-2 blot density/  $\beta$ -actin blot density for each individual. There was a significant reduction in PEN-2 expression in mutant fibroblasts versus control fibroblasts (\*\*\*) denotes  $p < 0.001$ , unpaired t-test).

## **4.5 The effect of *NCSTN* c.1125+1 G>A and *PSENEN* c.66\_67insG mutations on gamma-secretase enzyme activity *in vitro***

Globally, mutations have now been reported in three of the four components of the gamma-secretase complex in HS (*PSEN1*, *NCSTN* and *PSENEN*), firmly implicating gamma-secretase in disease pathogenesis. It was therefore hypothesised that the identified mutations would affect gamma-secretase enzyme activity.

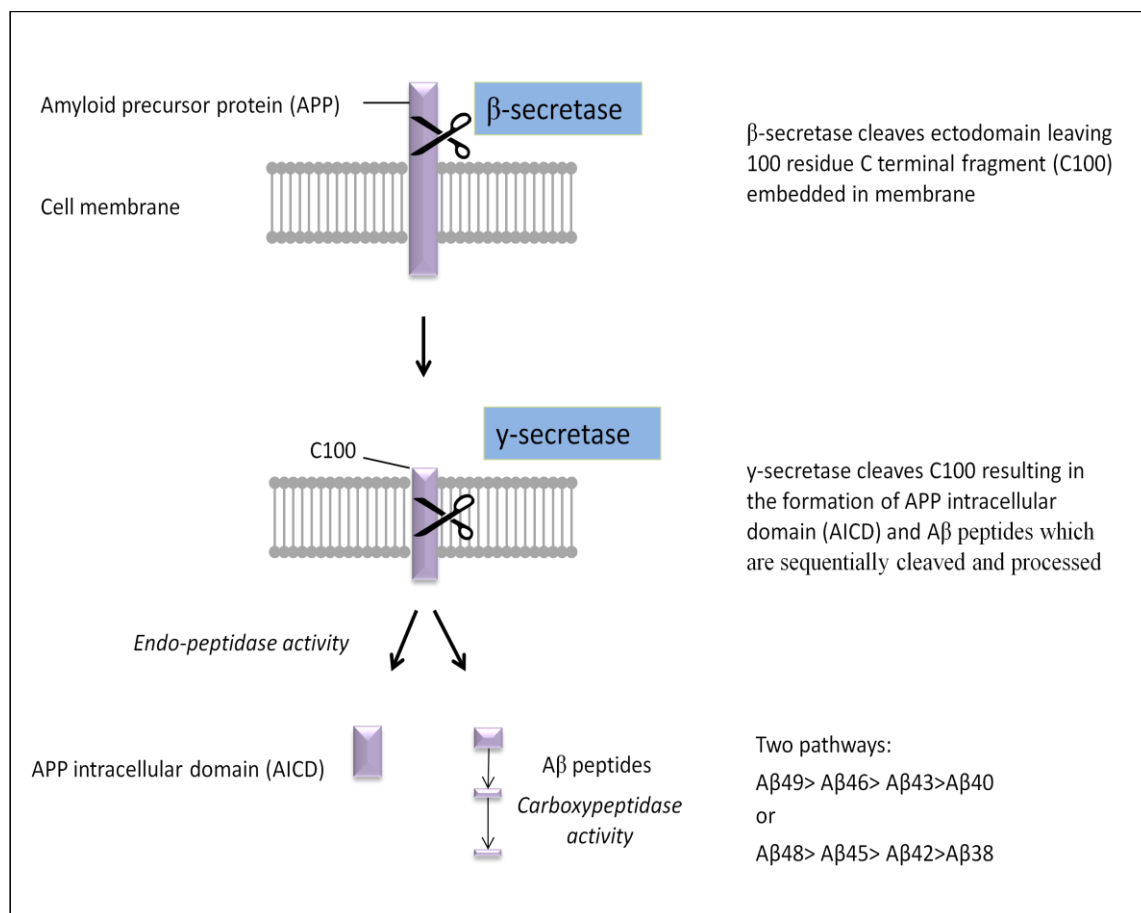
### **4.5.1 Methodological approaches to evaluating gamma-secretase activity *in vitro*.**

Multiple methods have been devised to assess gamma-secretase enzyme activity *in vitro*. The original and most commonly reported assay uses a Flag-tagged portion of APP protein (C100-Flag) as an enzyme substrate<sup>264; 265</sup>. The amount of Flag-tagged AICD that is produced reflects endopeptidase enzyme activity and the amount and relative ratio of A $\beta$  peptides produced reflects carboxypeptidase enzyme activity. Flag-tagged Notch is an alternative enzyme substrate with NICD-Flag production acting as a measure of endopeptidase activity. Carboxypeptidase activity is not as easy to assess with Notch based assays however as whilst N $\beta$  peptide levels can be measured, interpretation of the results remains at a premature stage<sup>178; 247</sup>. One commercial *in vitro* gamma-secretase activity assay was produced that utilised APP as the enzyme substrate however this lacked formal validation and was withdrawn from the market in 2010<sup>275</sup>. The original APP based method was deemed to be the most well-established, reliable, interpretable and informative method to determine gamma-secretase enzyme activity *in vitro* in this study.

### **4.5.2 Analysis of *in vitro* gamma-secretase activity using an amyloid precursor protein-based assay (C100F substrate)**

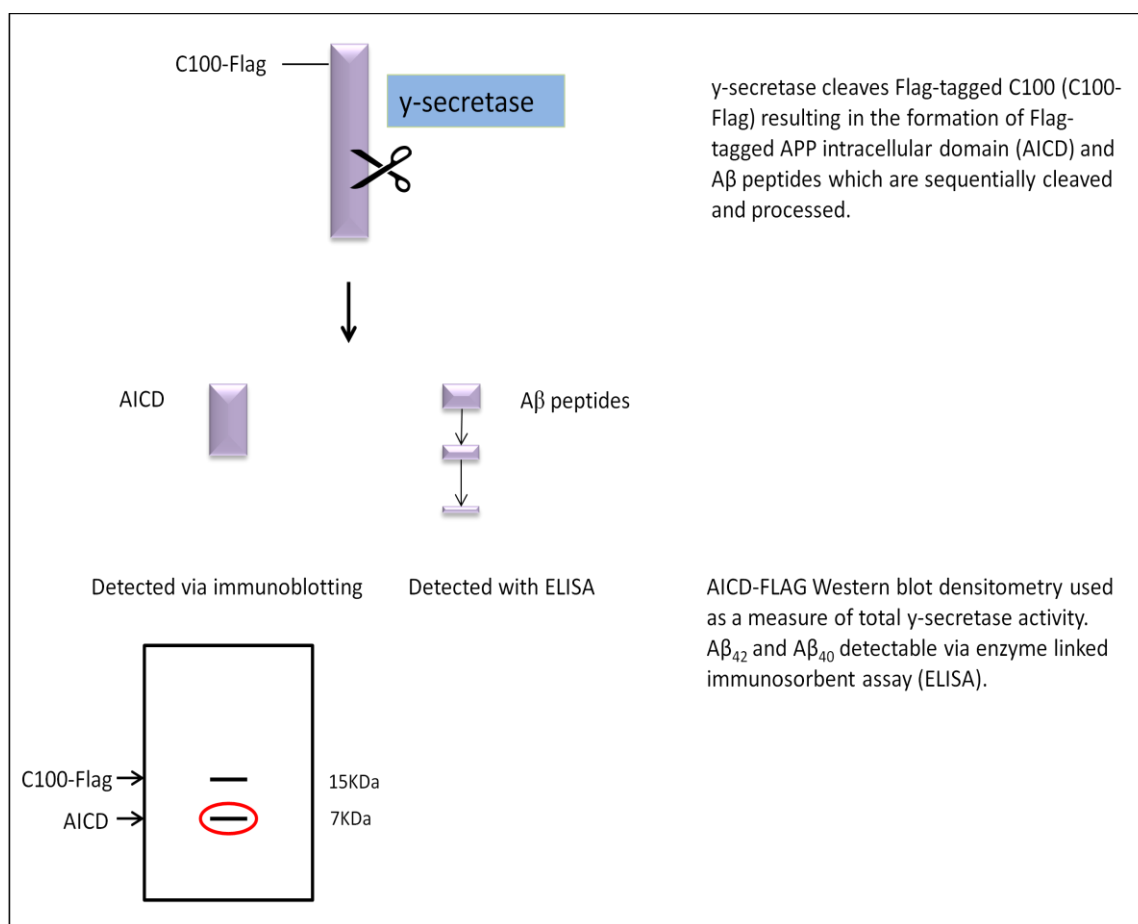
Amyloid precursor protein (APP) is a cell-membrane protein. The extra-cellular portion of the protein can be cleaved by either  $\alpha$ -secretase or  $\beta$ -secretase.  $\beta$ -secretase cleavage

leaves a 100 amino acid segment of APP (C100) within the cell membrane which in turn is cleaved by gamma-secretase to produce APP intracellular domain (AICD) and amyloid  $\beta$  peptides ( $A\beta$ ) ( $\epsilon$ -cleavage, endopeptidase activity). Two different length  $A\beta$  peptides can be produced following gamma-secretase  $\epsilon$ -cleavage,  $A\beta_{49}$  and  $A\beta_{48}$ , which are progressively processed by gamma-secretase and shortened (carboxypeptidase-like activity) as shown in Figure 4.15. It is this processing activity of gamma-secretase that is felt to be important in FAD, when the ratio of  $A\beta_{42}$ :  $A\beta_{40}$  is raised, potentially increasing the likelihood of  $A\beta$  deposition in the brain<sup>178</sup>.



**Figure 4.15.  $\beta$ -secretase and  $\gamma$ -secretase cleavage and processing of amyloid precursor protein in the cell membrane.**  $\beta$ -secretase cleaves APP to leave a 100aa C terminal fragment (C100) in the cell membrane. Gamma-secretase then cleaves that to form AICD and  $A\beta_{48}$  or  $A\beta_{49}$  peptides ( $\epsilon$ -cleavage, endopeptidase activity). The  $A\beta$  peptides are subsequently sequentially cleaved to produce shorter length peptides (carboxypeptidase-like activity).

The basis of the gamma-secretase assay employed here is that a Flag-tagged C100 peptide is produced and used as the enzyme substrate. In the correct environment, gamma-secretase (in the form of purified and solubilised membrane protein harvested from patient fibroblasts) cleaves this substrate to produce AICD-Flag and A $\beta$  peptides. AICD-Flag abundance is measured by immunoblotting (using an anti-Flag antibody) and that correlates with overall enzyme activity ( $\epsilon$ -cleavage). A $\beta$ 40 and A $\beta$ 42 are quantified by ELISA and the ratio of these two products is an indication of the processing (carboxypeptidase-like) activity of gamma-secretase (Figure 4.16).

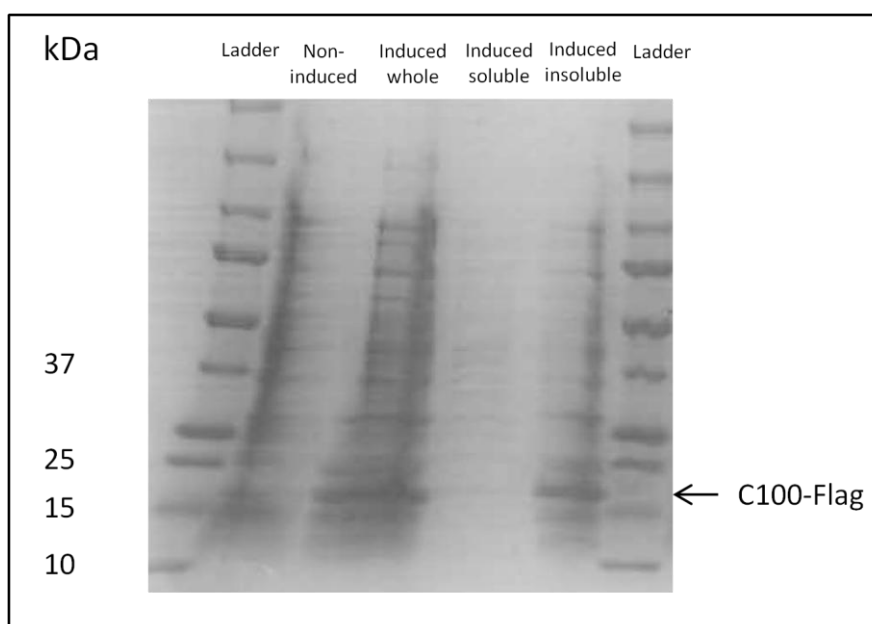


**Figure 4.16. Principle of the C100-Flag APP based gamma-secretase activity assay.** A C100-Flag substrate is produced which is cleaved by gamma-secretase to produce AICD-Flag (measure of endopeptidase activity) and A $\beta$  peptides (the ratio of A $\beta$ 42:A $\beta$ 40 is a measure of carboxypeptidase activity). AICD-Flag and A $\beta$  peptides can be quantified by immunoblotting and ELISA respectively.



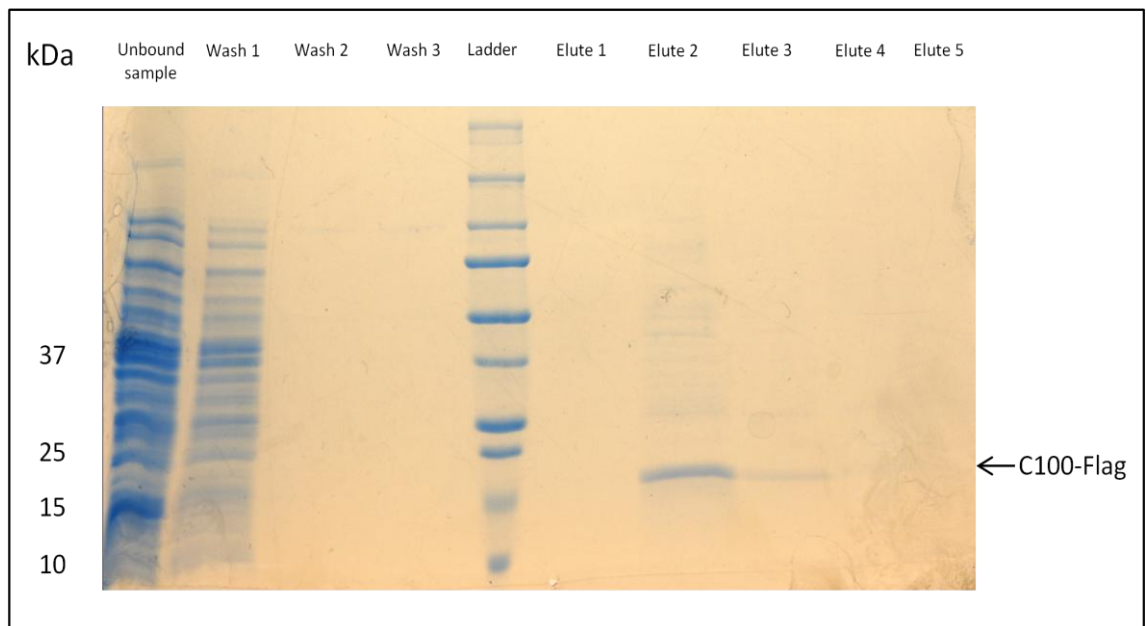
### **4.5.3 Production of the C100-Flag enzyme substrate**

A C100-Flag pET-21b (-) plasmid was kindly provided by Dr Oliver Holmes, Centre for Neurological Diseases, Harvard Medical School, Boston, USA<sup>264</sup>. C100-Flag protein was generated and purified from this plasmid following a protocol developed by the Centre for Neurological Diseases, Harvard Medical School. NEB 5-alpha competent e-coli were transformed with the C100-Flag pET-21b and cultured. The plasmid was subsequently purified generating a plasmid concentration of 330ng/μl. BL21 (DE3) competent E Coli were then transformed with the purified C100-Flag pET-21b plasmid and cultured. C100-Flag protein expression was induced with 1mM IPTG and cells were physically lysed in Triton X-100 using a glass homogeniser. The sample was spun to create a soluble and insoluble fraction. The insoluble fraction (cell pellet) contained cell membranes enriched with C100-Flag, the soluble fraction was discarded. Small aliquots of pre-IPTG induced lysed cells, post-IPTG induced lysed cells and the soluble and insoluble fraction of that lysate were analysed via electrophoresis to assess whether C100-Flag protein expression had indeed been induced. Staining of the protein blot with InstantBlue (Expedeon, Harston, UK) confirmed the presence of a protein band at the expected weight (15kDa) in samples from post-IPTG induced lysed cells, more specifically the insoluble fraction (Figure 4.17).



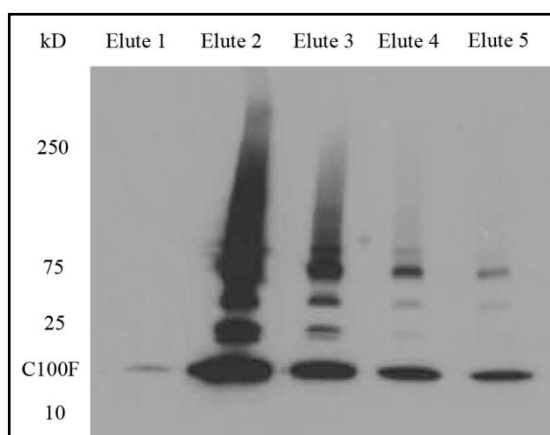
**Figure 4.17. Protein electrophoresis of pre- and post-IPTG induced cell lysate samples from BL21 (DE3) E.coli cells expressing the C100-Flag pET-21b plasmid.** A protein band can be seen at around 15kDa in the IPTG induced cell lysate consistent with C100-Flag production. Upon splitting this lysate in to soluble and insoluble fractions the protein can be seen in the insoluble fraction, as would be expected given the membranous location of the C100 protein.

The insoluble fraction (cell pellet) was re-suspended and homogenised with a glass homogeniser and C100-Flag protein was selected and purified using anti-flag tagged beads. The beads were washed three times and the C100-Flag protein was eluted from the beads with five separate aliquots of elution buffer. Aliquots of unbound sample, washes 1 – 3 and elutes 1 – 5 were analysed via electrophoresis and InstantBlue (Expedeon) staining. C100-Flag was present in elutes 2 and 3 (most abundant in elute 2, Figure 4.18).



**Figure 4.18. Electrophoresis and InstantBlue staining of samples generated during the purification of C100-Flag.** A protein band consistent with C100-Flag can be seen at around 15kDa in elutes 2 and 3.

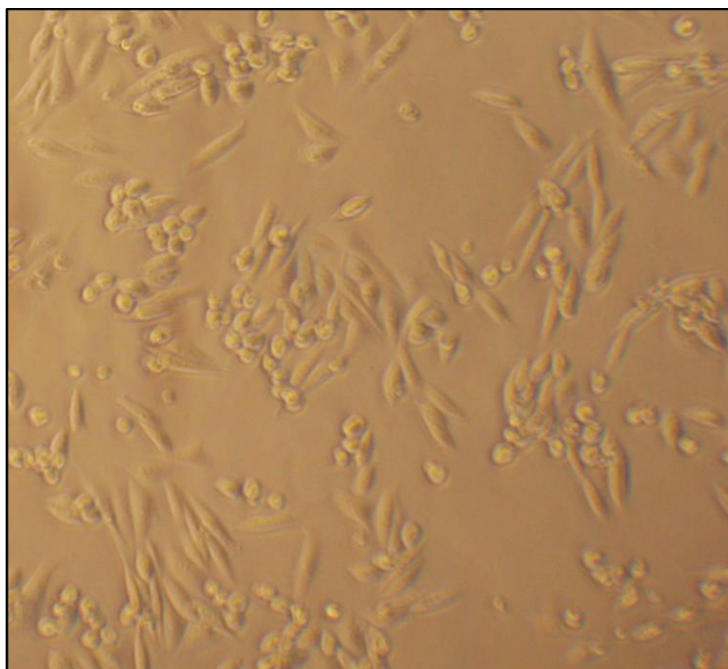
The five elute samples were assessed by immunoblotting with an anti-Flag primary antibody. C100-Flag was present in all five elutes, more so in elutes 2 and 3 (Figure 4.19). Elute 2 was therefore used as the source of C100-Flag in all gamma-secretase assays.



**Figure 4.19. Immunoblot of purified C100-Flag samples using a primary anti-Flag antibody.** A protein band can be seen at 15kDa in elutes 1 – 5 consistent with C100-Flag.

#### **4.5.4 Initiation and maintenance of an S-20 cell line**

S-20 is a Chinese hamster ovary (CHO) cell line that stably expresses Nicastrin-V5/His, human PSEN1, N-terminally FLAG-tagged Pen-2 and C-terminally HA-tagged Aph-1 $\alpha$ <sup>266</sup> (Figure 4.20). The relative over-expression of these components means that solubilised membrane protein harvested from these cells can be used as a positive control in gamma-secretase enzyme activity assays<sup>264; 266</sup>. S-20 cells were kindly provided by Dr Oliver Holmes, Centre for Neurological Diseases, Harvard Medical School, USA.



**Figure 4.20. S20 cells in culture.**

#### **4.5.5 Purification and solubilisation of membrane proteins from primary human fibroblasts and S20 cells**

Axillary skin was taken from the two patients harbouring mutations (affected skin) and three age-matched volunteer subjects (controls) with no history of HS. Primary human fibroblasts were isolated and cultured and the total number of cells harvested from each individual is shown in Table 4.2. S20 cells were cultured in parallel (Table 4.2).

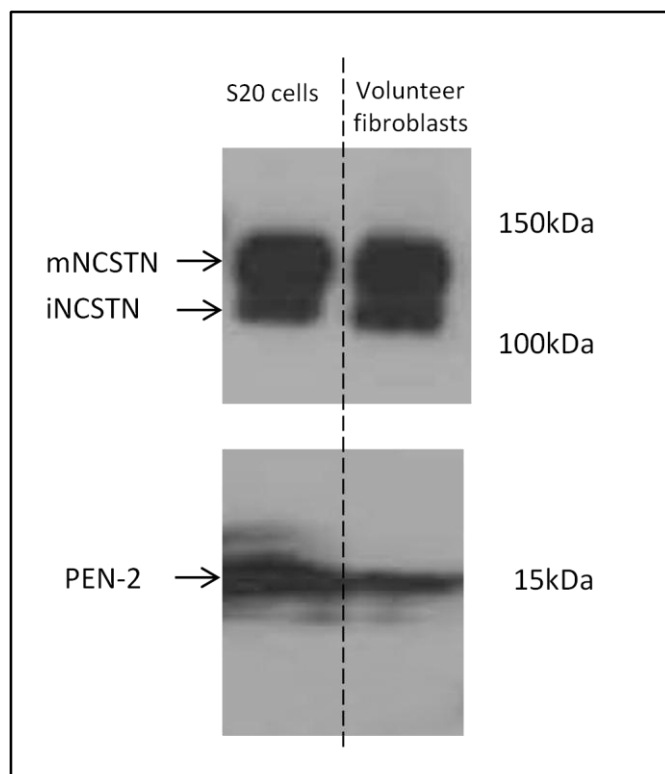
Membrane proteins were purified and solubilised using a protocol devised by our collaborators at Harvard Medical School, USA<sup>266</sup>. Cells were physically lysed and cell membranes were isolated and washed with sequential ultra-centrifugation and sodium bicarbonate. Membranes were solubilised in 1% CHAPSO (amount of 1% CHAPSO added shown in Table 4.2) and the sample was further centrifuged. The supernatant was defined as solubilised gamma-secretase preparation and was used as crude gamma-secretase fractions in gamma-secretase assays. The number of fibroblasts used to produce each microlitre of solubilised membrane protein was based on our collaborators protocol<sup>266</sup>.

**Table 4.2. Number of cells harvested and amount of 1% CHAPSO used to extract solubilised membrane proteins**

Sample	Number of primary human fibroblasts harvested	Amount of 1% CHAPSO added/ $\mu$ l
S20	$5.475 \times 10^8$	400
<i>PSENEN</i> mutant	$4.225 \times 10^7$	140
<i>NCSTN</i> mutant	$3.95 \times 10^7$	140
Control 1	$4.1 \times 10^7$	140
Control 2	$4.025 \times 10^7$	140
Control 3	$3.825 \times 10^7$	140

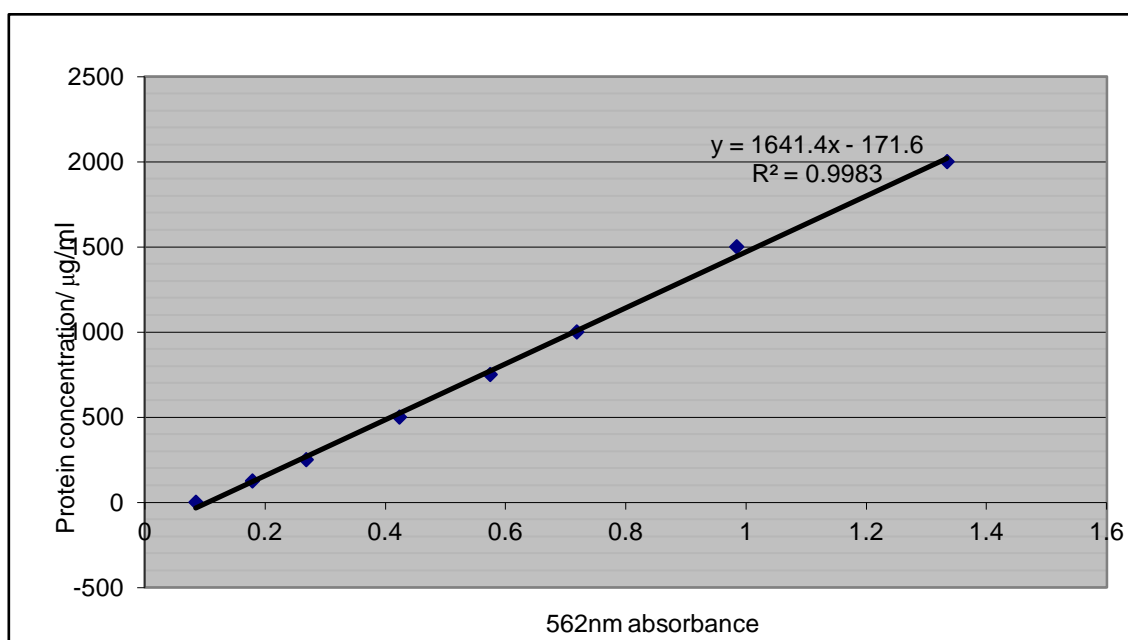
Solubilised gamma-secretase preparations derived from S20 cells and primary human fibroblasts (control 1) were subjected to electrophoresis and probed with anti-nicastrin and anti-PEN-2 antibodies. Nicastrin and PEN-2 were detected in both samples, thus it

was assumed that the gamma-secretase components had been effectively solubilised (Figure 4.21).



**Figure 4.21. Immunoblotting of NCSTN and PEN-2 in solubilised membrane preparations from S20 cells and primary human fibroblasts.** There was good expression of both proteins in solubilised membrane preparations harvested from S20 and primary human fibroblasts.

Solubilised gamma-secretase preparation concentrations were measured using a BCA protein concentration assay. All assays were performed in triplicate and the standard curve for the albumin controls is shown in Figure 4.22. Protein concentrations for all of the solubilised membrane protein samples were calculated by plotting light absorbance readings against this curve. Protein concentrations are shown in Table 4.3.



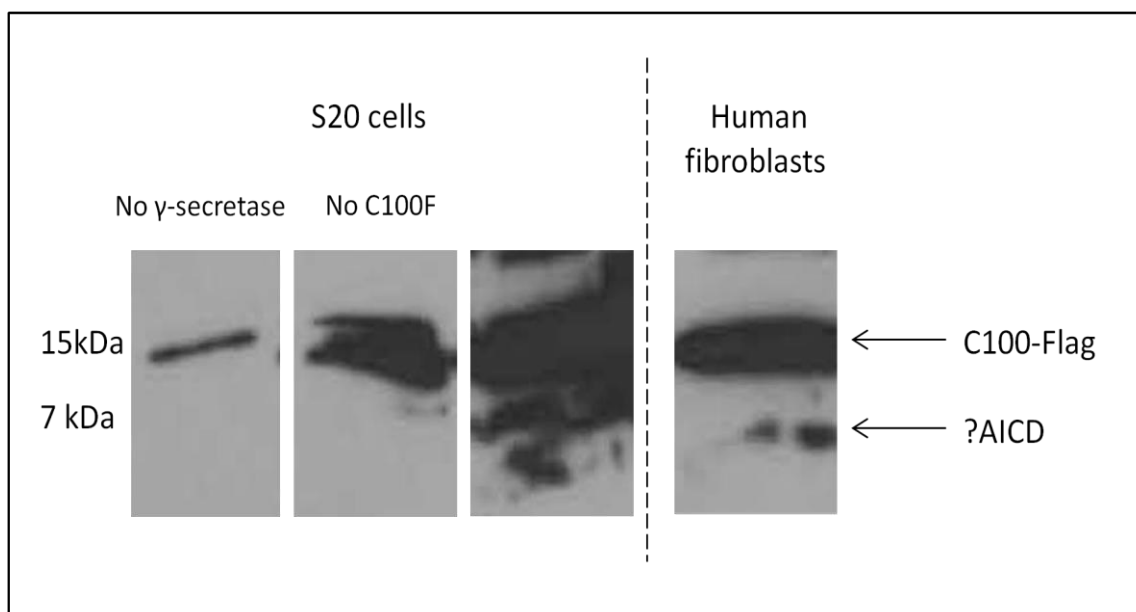
**Figure 4.22. Standard curve plotted using protein concentration and 562nm light absorbance readings for protein standards.** The protein concentration of solubilised cell membrane preparations were calculated from their 562nm light absorbance readings using this standard curve.

**Table 4.3. Concentration of the solubilised cell membrane preparations as calculated using BCA assays.**

Sample	Solubilised membrane protein concentration/ µg/ ml
S20	1624.3
PSENEN mutant	565.6
NCSTN mutant	943.1
Control 1	737.9
Control 2	1780.2
Control 3	409.6

#### 4.5.6 Optimisation of the gamma-secretase activity assay

Gamma-secretase assays were performed according to a protocol devised by our collaborators at Harvard Medical School<sup>264</sup>. 8µg solubilised gamma-secretase preparations harvested from S20 cells and primary human fibroblasts were incubated with cholesterol, L-α-phosphatidylcholine and L-α-phosphatidylethanolamine. The lipids form detergent-lipid mixed micelles which provide a platform upon which gamma-secretase can cleave its substrates. C100-Flag substrate was then added and the mixture was left to incubate for 4 hours at 37°C. Runs with no gamma-secretase (ie. no solubilised gamma-secretase preparation added) and no substrate were performed as negative control experiments. 12µl of the reaction mixtures were subjected to electrophoresis and probed with an anti-Flag antibody (Figure 4.23).



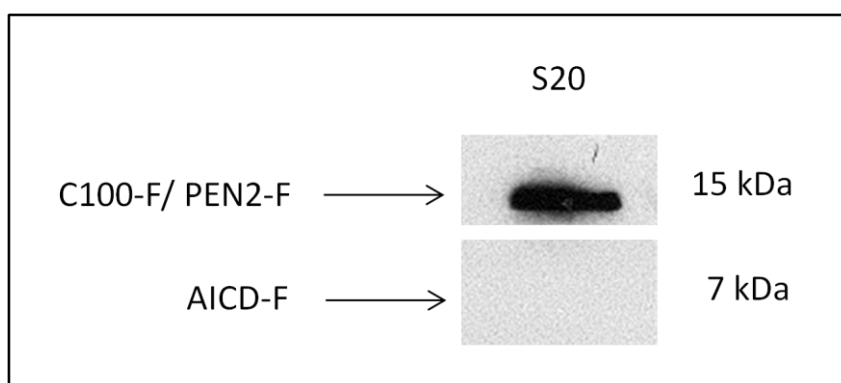
**Figure 4.23. Immunoblot showing the results of a gamma-secretase activity assay.** The C100-Flag substrate runs at 15kDa, the expected enzyme product, AICD, is predicted to run at 7kDa. The C100-Flag substrate can be seen at 15kDa (the band seen in the no C100F lane represents PEN2-Flag expressed by S20 cells). AICD appeared to be visible in reaction mixes comprising gamma-secretase harvested from S20 cells and primary human fibroblasts but not visible in the no gamma-secretase/ no substrate runs, as would be expected.



A 15kDa band could be seen in the no gamma-secretase run, consistent with the C100-Flag substrate. A 15kDa band could also be seen in the no substrate reaction. As the solubilised gamma-secretase preparation used in this reaction was harvested from S20 cells, this band represents the Flag-tagged PEN-2 protein produced by those cells. A very intense signal was observed around 15kDa in the S20 reaction which is likely to represent both the Flag-tagged substrate and PEN-2 present in that reaction. Any lighter bands (AICD expected at 7kDa) were therefore hard to identify however there did appear to be a band visible at 7kDa. The C100-Flag band seen at 15kDa in the primary human fibroblast run was smaller and more well-defined than the band seen in the S20 run. There did appear to be a band visible at 7kDa however this was broken, possibly as a consequence of an air bubble affecting protein transfer on to the ECL membrane. Overall, there appeared to be some evidence that AICD was being produced in these activity assays, however the band expected at 7kDa was not clear and was vastly overshadowed by the substrate band seen at 15kDa in the S20 reaction. As a consequence, the protein blots were cut at 13kDa prior to exposure in subsequent runs.

#### **4.5.6.1 Optimisation step 1 – Cutting the ECL membrane at 13kDa prior to exposure to separate the substrate and AICD signals**

All optimisation experiments were designed in collaboration with Dr. Oliver Holmes, Harvard Medical School, USA. S20 cells were used in all optimisation experiments as they are known to over-express the four components of the gamma-secretase complex. The experiment was performed exactly as above however the ECL membrane was cut at 13kDa to form two segments, one theoretically containing C100-Flag (15kDa) and one AICD-Flag (7kDa). The immunoblot can be seen in Figure 4.24. Whilst the C100-Flag substrate/ Flag-tagged PEN-2 protein was clearly visible at 15kDa there was no signal observed at 7kDa suggestive of AICD production.

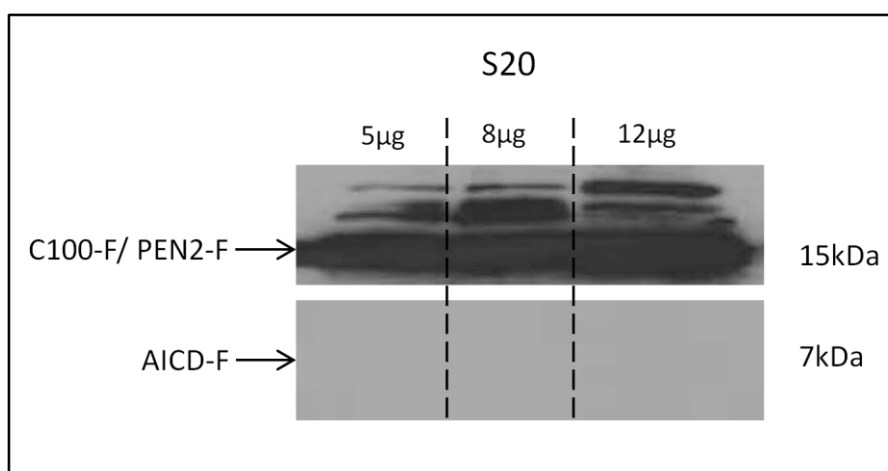


**Figure 4.24. Immunoblot of a gamma-secretase activity assay reaction performed using enzyme derived from S20 cells.** The C100-Flag substrate was visible at 15kDa but no AICD product was detectable.

Films were left for over 60 minutes and no bands were observed. On the assumption that there may only be a very small amount of AICD produced, the amount of protein blotted was increased from 12µl to 30µl. No band was observed at 7kDa with this increased protein load. ECL membranes were cut at 13kDa and 30µl protein was blotted in all future optimisation experiments.

#### **4.5.6.2 Optimisation step 2 – Altering the amount of solubilised gamma-secretase preparation added to each assay**

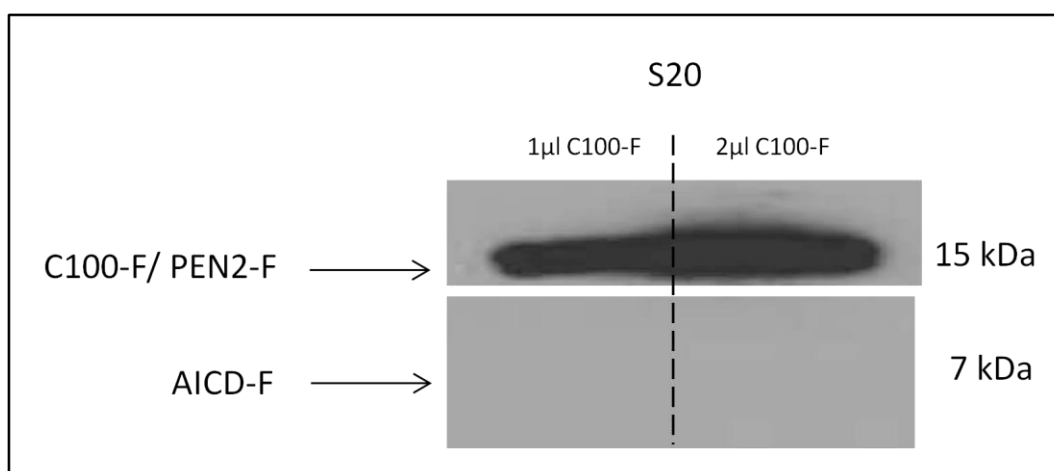
To ascertain the impact of altering the amount of solubilised cell membrane protein added to each reaction 5µg, 8µg and 12µg protein were added to three different reactions. This was done in the knowledge that higher concentrations would increase the final CHAPSO concentration in the reaction above the optimal 0.25%. CHAPSO concentrations above 0.45% have been shown to reduce enzyme activity, possibly by affecting micelle formation (Holmes, unpublished data). 30µl of reaction product was subjected to electrophoresis and probed with anti-Flag antibody. The C100-Flag/ PEN-2-Flag band at 15kDa was visible however no AICD-Flag product was evident on prolonged exposure (Figure 4.25).



**Figure 4.25. Immunoblot of enzyme activity assays using variable concentrations of S20 solubilised membrane protein.** The C100-Flag substrate is visible however no AICD-Flag was produced.

#### **4.5.6.3 Optimisation step 3 – Altering the amount of C100-F substrate added to each assay**

The C100-Flag substrate was purified effectively (Figure 4.19) and was visible in enzyme activity assays in which no solubilised proteins were added (Figure 4.23). Harvard use 1µl of purified substrate and this amount is sufficient to saturate the gamma-secretase complexes that are present. This means that the amount of substrate available is not a variable in activity assay outcomes. To ensure that assays were not failing due to a lack of substrate the amount of C100-Flag added was doubled to 2µg. Again the C100-Flag/ PEN-2-Flag band was evident at 15kDa but no AICD product was produced (see Figure 4.26).



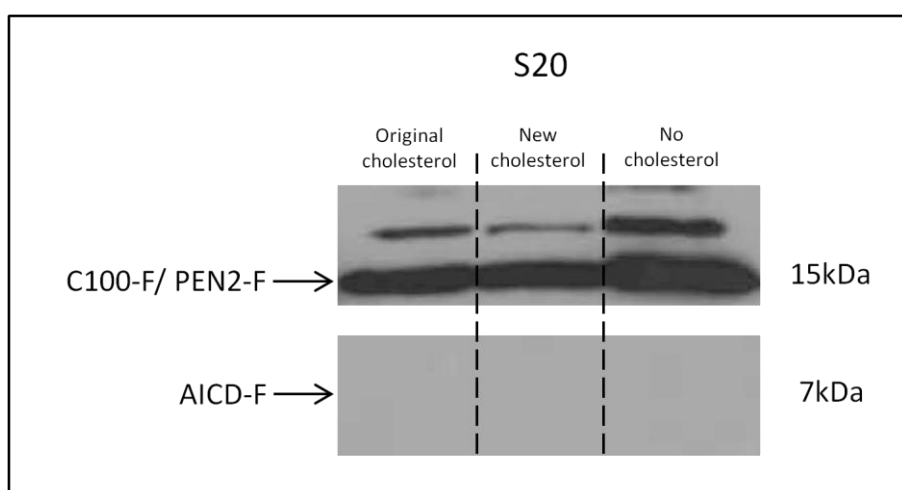
**Figure 4.26. Immunoblot of gamma-secretase assays using variable amounts of C100-Flag substrate.** No AICD-Flag product can be seen in reactions involving 1 or 2µl C100-Flag.

In summary, immunoblotting of the solubilised gamma-secretase preparations produced from S20 cells and human fibroblasts revealed the presence of the gamma-secretase components NCSTN and PEN-2 (Figure 4.21), inferring that membrane proteins had been isolated and solubilised effectively. The C100-Flag substrate was purified effectively (Figure 4.19) and was clearly visible in enzyme activity assays in which no solubilised proteins were added (Figure 4.23). The quantity added to each reaction was as stated in our collaborators protocol and doubling that amount made no impact on the efficacy of the assay. Gamma-secretase complexes and the C100-Flag substrate would both therefore appear to be present, thus it was assumed that the enzyme activity assay itself was not working effectively. If there was indeed any AICD produced in the very first optimisation run, which is unclear from the immunoblot performed, that would infer that something had changed in subsequent assays. The method, techniques and machinery used did not alter between runs. Reagent batches did not change and storage and handling of each reagent was as advised by the supplier. Any degradation or conformational/oxidative changes affecting the lipids could have suppressed micelle formation. These micelles are integral to the assay as they support the gamma-secretase

complexes and facilitate enzyme activity. A further consideration was that there can be significant batch to batch variation in cholesterol quality. Our collaborators have previously shown that some samples even impair enzyme activity (Holmes, unpublished data). A new batch of cholesterol was therefore used for the next assay and a no cholesterol run was performed.

#### **4.5.6.4 Optimisation step 4 – Cholesterol batch changed and a no-cholesterol assay performed**

The assay was repeated three times with 8µg solubilised gamma-secretase preparation. Cholesterol from a new batch was used in one run, cholesterol from the previous experiments was used in a second run and no cholesterol was added to the final reaction. None of these interventions conferred a benefit. No AICD-Flag was detected in any of the three runs (Figure 4.27).



**Figure 4.27. Immunoblot of gamma-secretase enzyme activity assays using different batches of and no cholesterol.** The C100-Flag/ PEN-2-Flag band can be seen at 15kDa but no AICD-Flag was evident in any of the three runs.

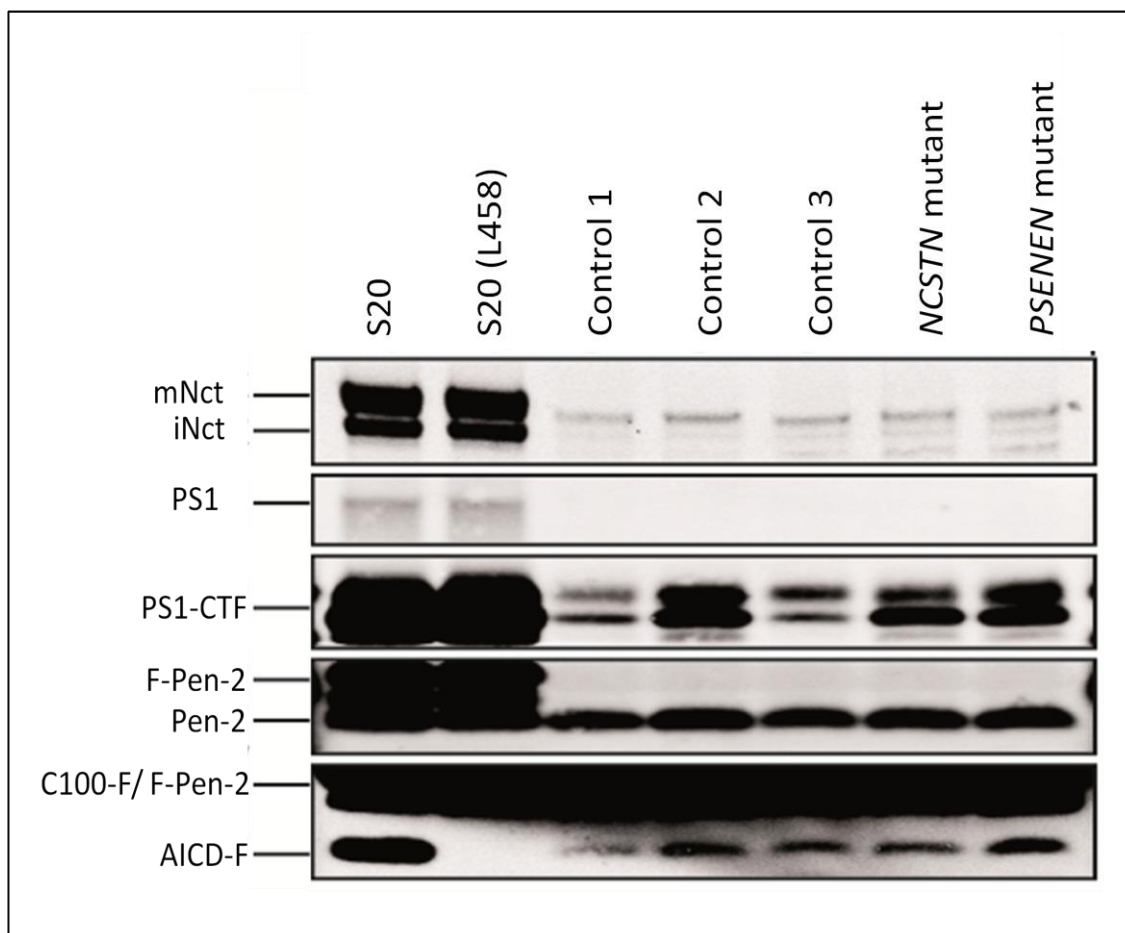
#### **4.5.6.5 Enzyme activity assays were performed at Harvard Medical School**

From the available data, it would appear that both the enzyme and substrate were present in these assays but that there was a problem with the enzyme activity assay itself. Despite the above optimisation steps the nature of that problem could not be identified and therefore corrected. Our collaborators, based at The Centre for Neurological Diseases, Harvard Medical School, USA, have optimised and routinely perform this assay in their laboratory. It was therefore decided that the solubilised gamma-secretase samples and C100-Flag substrate should be sent to their laboratory for analysis. All activity assays were performed with the solubilised gamma-secretase samples and C100-Flag substrate generated in our laboratory. Experiments detailed below were designed in collaboration with Dr Oliver Holmes and performed by Dr Oliver Holmes (Harvard Medical School, Centre for Neurological Diseases, USA).

#### **4.5.7 Gamma-secretase activity assays**

The quantity of solubilised gamma-secretase preparation added to each reaction was standardised at 8µg. All activity assays were repeated 7 times for each of the six samples (harvested from S20 cells, *NCSTN* mutant fibroblasts, *PSENEN* mutant fibroblasts and control fibroblasts 1-3) to ensure consistent results. Assay reactions were analysed by electrophoresis and immunoblotting with primary antibodies to NCSTN, PEN-2, PSEN1, PSEN1-CTF and Flag. A blot from one activity assay is shown in Figure 4.28. Bands corresponding to the gamma-secretase components, C100-Flag substrate and AICD-Flag product are clearly visible implying that the solubilised gamma-secretase preparations and C100-F substrate prepared in-house were of sufficient quality and concentration. S20 cells were included as a positive control in this study and, as would be expected, there was an over-expression of gamma-secretase components and a correspondingly high amount of AICD-Flag produced. This therefore acted as a suitable control against which to compare AICD production in the other

samples. To confirm that the AICD bands visualised represented products of gamma-secretase cleavage an assay was run with the gamma-secretase inhibitor L458 (negative control). This successfully prevented any AICD-Flag production, confirming the specificity of the assay.



**Figure 4.28 A representative gamma-secretase activity assay immunoblot.** Bands can be seen representing three of the gamma-secretase components NCSTN (mature, mNct, and immature, iNct), PSEN1 (PS1, PS1-CTF), PEN-2 (Pen2-Flag, Pen-2), the substrate (C100-Flag) and the enzyme product (AICD-Flag).

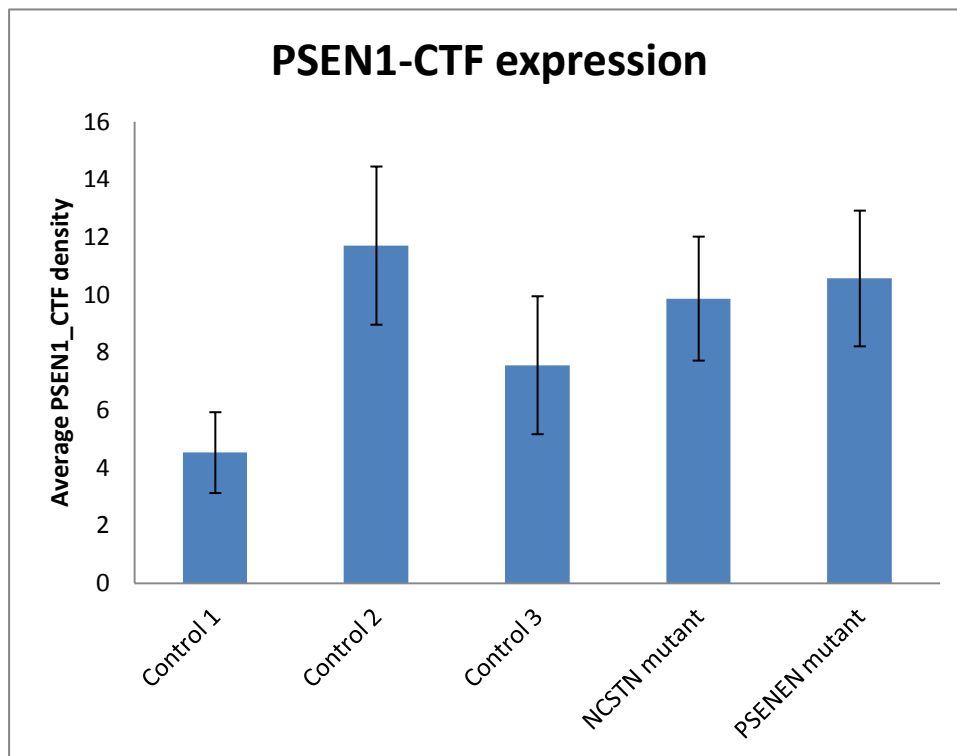
#### **4.5.8 Analysis of gamma-secretase enzyme activity in mutant and control primary human fibroblasts**

##### **4.5.8.1 Hypothesis 1: There will be fewer gamma-secretase complexes and a consequent reduction in enzyme activity in samples harvested from mutant fibroblasts versus control fibroblasts**

One might hypothesise that a reduced expression of any one component of the gamma-secretase complex (as would appear to be the case with the *NCSTN* and *PSENEN* mutations reported here) would impact upon the number of mature gamma-secretase complexes present at the level of the cell membrane which may consequently alter enzyme activity at that site.

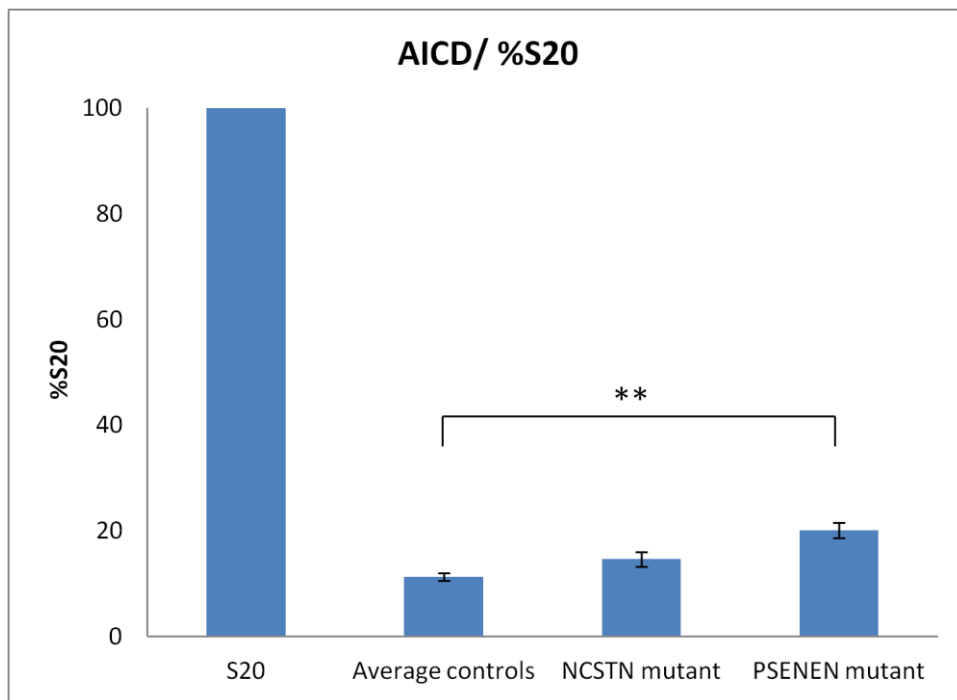
PSEN1-CTF is a stable marker of the number of mature complexes present in any sample as it is only produced following successful cleavage of PSEN1 (required for complex activity). Whilst PSEN1-CTF levels could be measured and compared across individuals in this experiment, there was no reliable ubiquitously expressed endogenous control that could be used to confirm equal protein loading on immunoblotting (similar to  $\beta$ -actin used in earlier immunoblotting experiments). The original amount of protein added to each reaction is however standardised, so excluding any manual handling errors (eg. pipetting error) then equal amounts of protein should be added to each reaction. The seven experimental repeats would also largely negate any such pipetting errors. The average PSEN1\_CTF expression for each sample is shown in Figure 4.29. Significant variability was observed between the control fibroblast samples but there appeared to be no significant difference in PSEN1\_CTF expression in mutant fibroblasts versus controls.





**Figure 4.29. PSEN1-CTF expression in solubilised membrane preparations harvested from mutant and control fibroblasts.** There was some variability in PSEN1-CTF expression observed between controls but there was no significant difference in PSEN1-CTF expression in the *NCSTN* or *PSENEN* mutant fibroblasts versus controls (error bars = standard error (SE)).

Crude gamma-secretase activity per mg solubilised membrane protein in each sample was analysed by comparing AICD-Flag production in every assay against the positive control (AICD-Flag production using solubilised membrane samples harvested from S20 cells). Average AICD-Flag values were calculated across all seven runs. No significant difference was detected between the *NCSTN* mutant and controls (Figure 4.30). If anything, a slightly increased enzyme activity was observed in *PSENEN* mutant fibroblasts vs. controls however the difference was small.

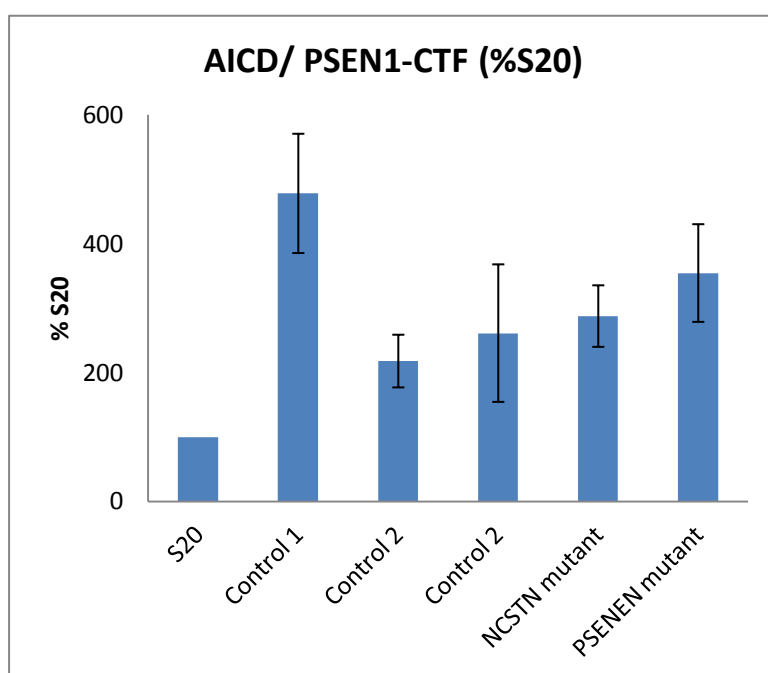


**Figure 4.30. Crude gamma-secretase activity per mg of total solubilised membrane protein expressed as a percentage of the positive control (S20 cells).** There was no significant difference in enzyme activity observed between the *NCSTN* mutant fibroblasts and controls. There was a small and significant increase in enzyme activity observed between *PSENEN* mutant fibroblasts and controls (error bars = SE, \*\* denotes  $p < 0.01$ , one-way ANOVA with Dunn's correction).

Overall, these data were not consistent with the hypothesis that there would be a reduction in complex number or crude gamma-secretase enzyme activity in mutant fibroblasts. PSEN-CTF expression was no different in mutant fibroblasts versus controls and crude enzyme activity appeared slightly increased, rather than reduced as was hypothesised, in *PSENEN* mutant fibroblasts. Crude gamma-secretase activity may be considered an indirect measure of mature gamma-secretase complex number if mature complexes are assumed to be of equal activity in all individuals. The indifferent PSEN1-CTF expression but apparent increase in enzyme activity observed in *PSENEN* mutant fibroblasts may however infer that enzyme activity per complex may differ between individuals.

#### 4.5.8.2 Hypothesis 2: Gamma-secretase enzyme activity per complex will be reduced in samples harvested from mutant fibroblasts versus control fibroblasts

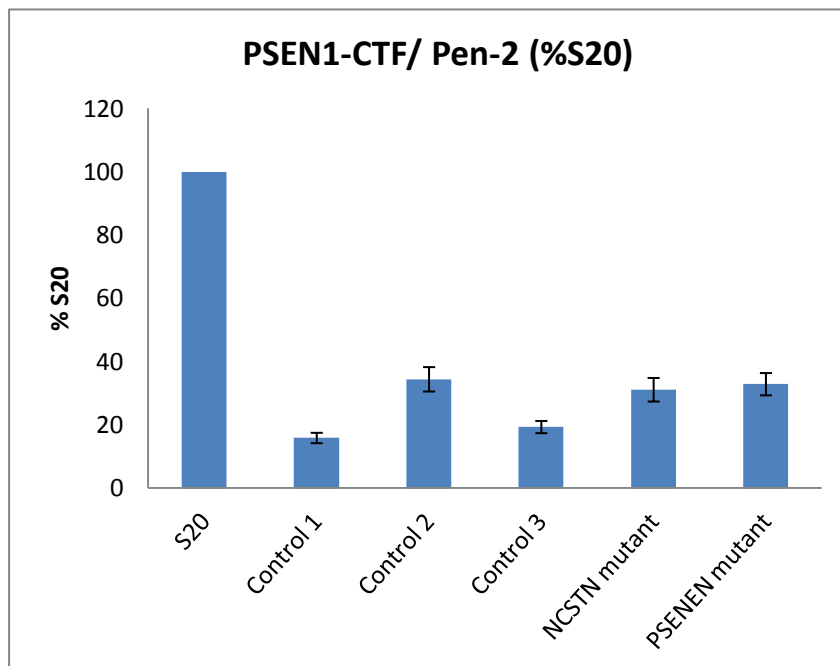
To measure enzyme activity per gamma-secretase complex AICD-Flag was controlled against PSEN1-CTF, a marker of the number of mature complexes present. The AICD-Flag/ PSEN1-CTF ratio for each individual was expressed as a percentage of the value obtained for the positive control S20 sample. Averages were calculated across all seven runs and are shown in Figure 4.31. There was no significant difference in enzyme activity per complex in mutant fibroblasts vs. controls.



**Figure 4.31. Gamma-secretase enzyme activity per complex expressed as a percentage of the positive control (S20 cell line).** There was no significant difference in enzyme activity per complex between mutant fibroblasts and controls (error bars = SE).

#### **4.5.8.3 Hypothesis 3: Gamma-secretase complexes will demonstrate abnormal maturation in samples harvested from mutant fibroblasts versus controls**

As discussed in the introduction, the presence of all four gamma-secretase components is required for appropriate trafficking, maturation, stability and ultimately activity of the complex. It may therefore be hypothesised that a reduction in the available nicastrin or PEN-2 in mutation-positive patients, as would be suggested by earlier functional studies, may detrimentally affect complex maturity. A measure of complex maturity is the PSEN1-CTF/ PEN-2 ratio. One might imagine that this ratio would differ in the individual harbouring the *PSENEN* mutation for example because less PEN-2 may be available for uptake. Alternatively the PEN-2 incorporated in to the complexes may not facilitate effective proteolysis of PSEN1 which is necessary to produce PSEN1-CTF (required for catalytic activity). Ratios were calculated for all seven assays and the average result for each sample is presented in Figure 4.32. The ratio did not significantly differ in mutant fibroblasts vs. controls.

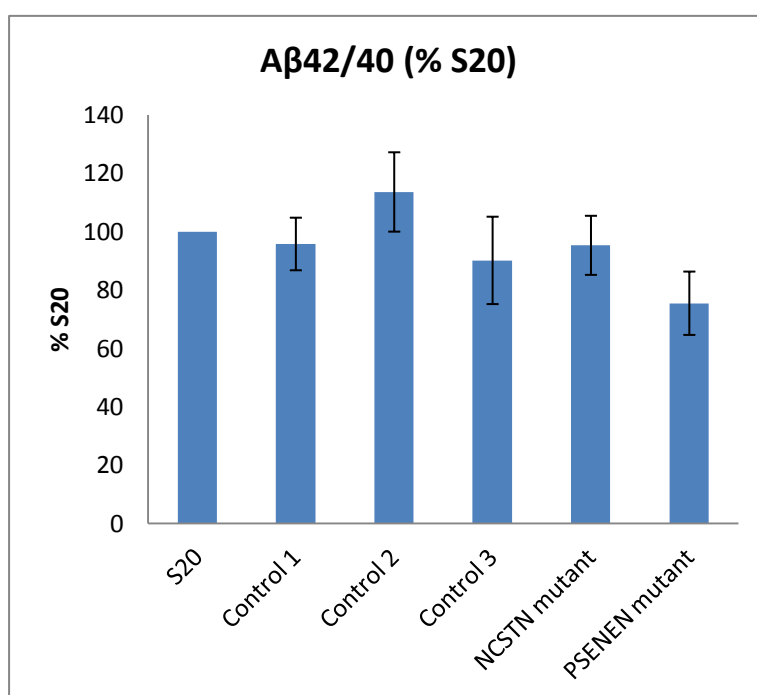


**Figure 4.32. PSEN1 CTF/ PEN-2 ratio for each sample expressed as a percentage of the ratio observed in the positive control S20 cells.** This ratio is representative of gamma-secretase complex maturity. There was no discernible difference in complex maturity in mutant fibroblasts versus control subjects (error bars = SE).

#### **4.5.8.4 Hypothesis 4: Gamma-secretase carboxypeptidase-like activity (processivity) will be altered in samples harvested from mutant fibroblasts vs. control fibroblasts.**

The enzyme activity data presented above relates to the  $\epsilon$ -cleavage or endopeptidase activity (production of AICD) of gamma-secretase which is regarded as a marker of overall enzyme activity. However, as stated in the introduction, the presenilin mutations reported in FAD unanimously affect carboxypeptidase activity but only rarely the endopeptidase activity of gamma-secretase. It is therefore conceivable that the nicastrin and presenilin mutations presented here may affect the carboxypeptidase-like processing function of gamma-secretase rather than the endopeptidase activity. ELISA's were performed to quantify the amount of A $\beta$ 42 and A $\beta$ 40 produced in each enzyme activity assay. A $\beta$ 42:A $\beta$ 40 ratios were calculated for every run and presented as a

fraction of the ratio derived in the positive control S20 reaction. The average ratio was calculated across all seven assays and the results are shown in Figure 4.33. There was no significant difference in the A $\beta$ 42:A $\beta$ 40 ratio in assays run with samples harvested from *NCSTN* mutant fibroblasts versus control fibroblasts. A slightly reduced ratio was observed in assays run with samples harvested from *PSENEN* mutant fibroblasts but this was not statistically significant.

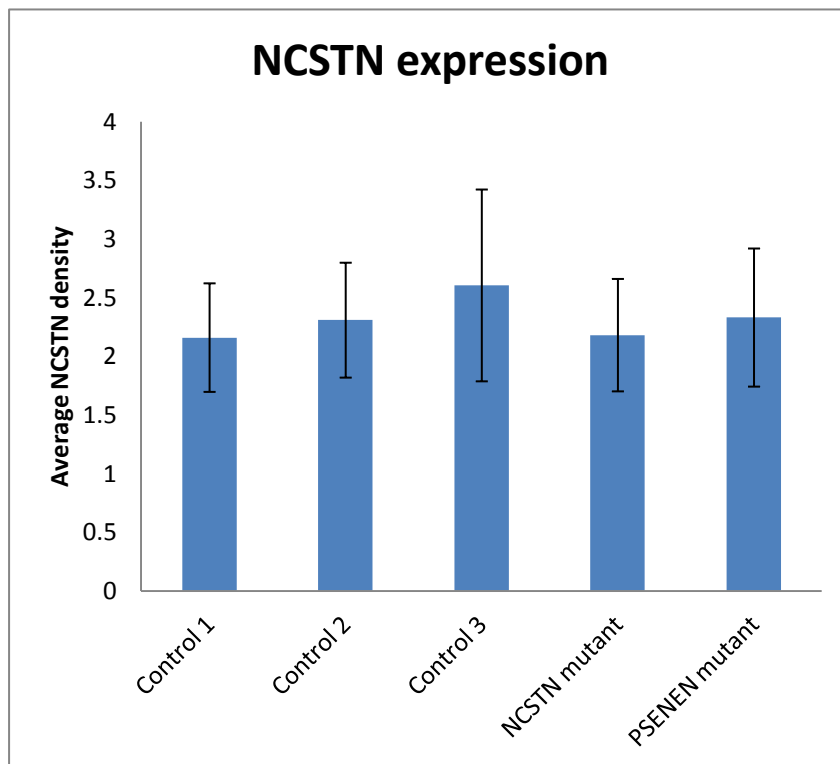


**Figure 4.33. A $\beta$ 42: A $\beta$ 40 ratios expressed as a percentage of the ratio derived from S20 cells.** There was no significant difference in the ratio derived in assays run with samples harvested from mutant fibroblasts versus control fibroblasts (error bars = SE).

#### **4.5.9 In contrast to total cell fractions, NCSTN and PEN-2 expression in solubilised cell membrane fractions is equal in mutant and control fibroblasts**

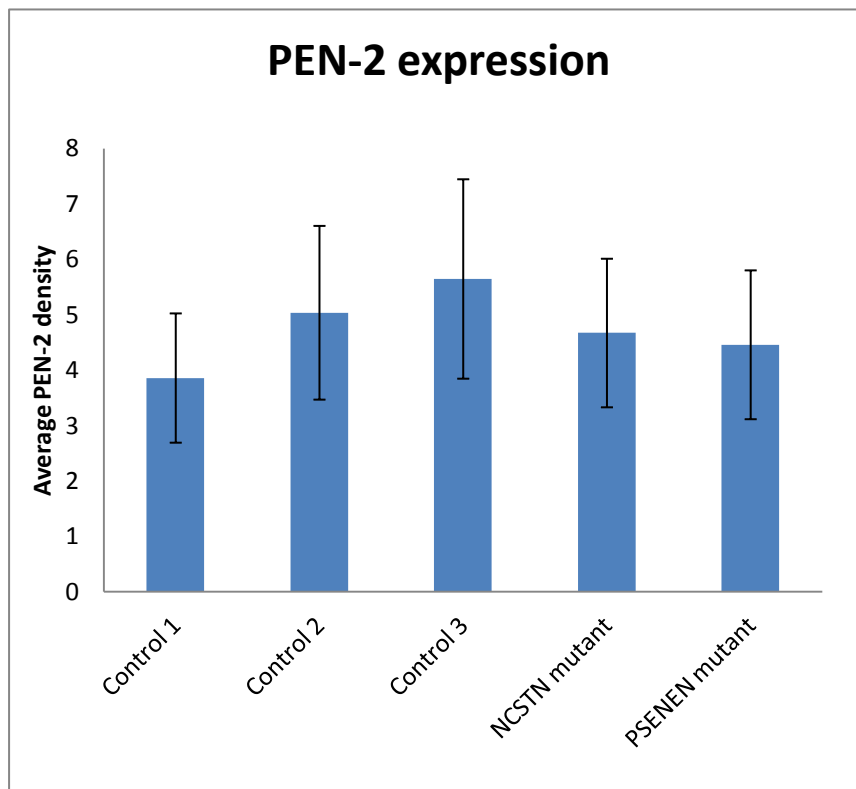
NCSTN and PEN-2 expression was reduced in total cell fractions harvested from *NCSTN* and *PSENEN* mutant fibroblasts respectively *in vitro*. (4.3.4, 4.4.4). Their relative expression in the solubilised cell membrane fractions harvested from the same

individuals for the purposes of the enzyme activity assays was therefore analysed as a comparison. Solubilised cell membrane preparations (also referred to as solubilised gamma-secretase preparation) were harvested from mutant and control fibroblasts for the purposes of the enzyme activity assays. NCSTN and PEN-2 blot densities were calculated from each of the seven independent enzyme activity assays and average densities (representing average protein expression) were calculated. It was very interesting to note that, in contrast to total cell fractions, there was no significant difference in NCSTN and PEN-2 expression in solubilised membrane fractions harvested from mutant and control fibroblasts *in vitro* (Figure 4.34, Figure 4.35). It should be noted however that there was no endogenous solubilised cell membrane protein that could be used as a reliable loading control for this analysis (as was the case when studying PSEN1-CTF expression in these same samples, section 4.5.8.1). All sample concentrations were however measured and equalised using BCA and the seven experimental repeats should have minimised any pipetting error.



**Figure 4.34. NCSTN expression in solubilised cell membrane fractions harvested from mutant and control fibroblasts.** Data represents the average protein band density measured across seven independent experiments. There was no significant difference in NCSTN expression in *NCSTN* mutant fibroblasts versus controls (error bars = SE).





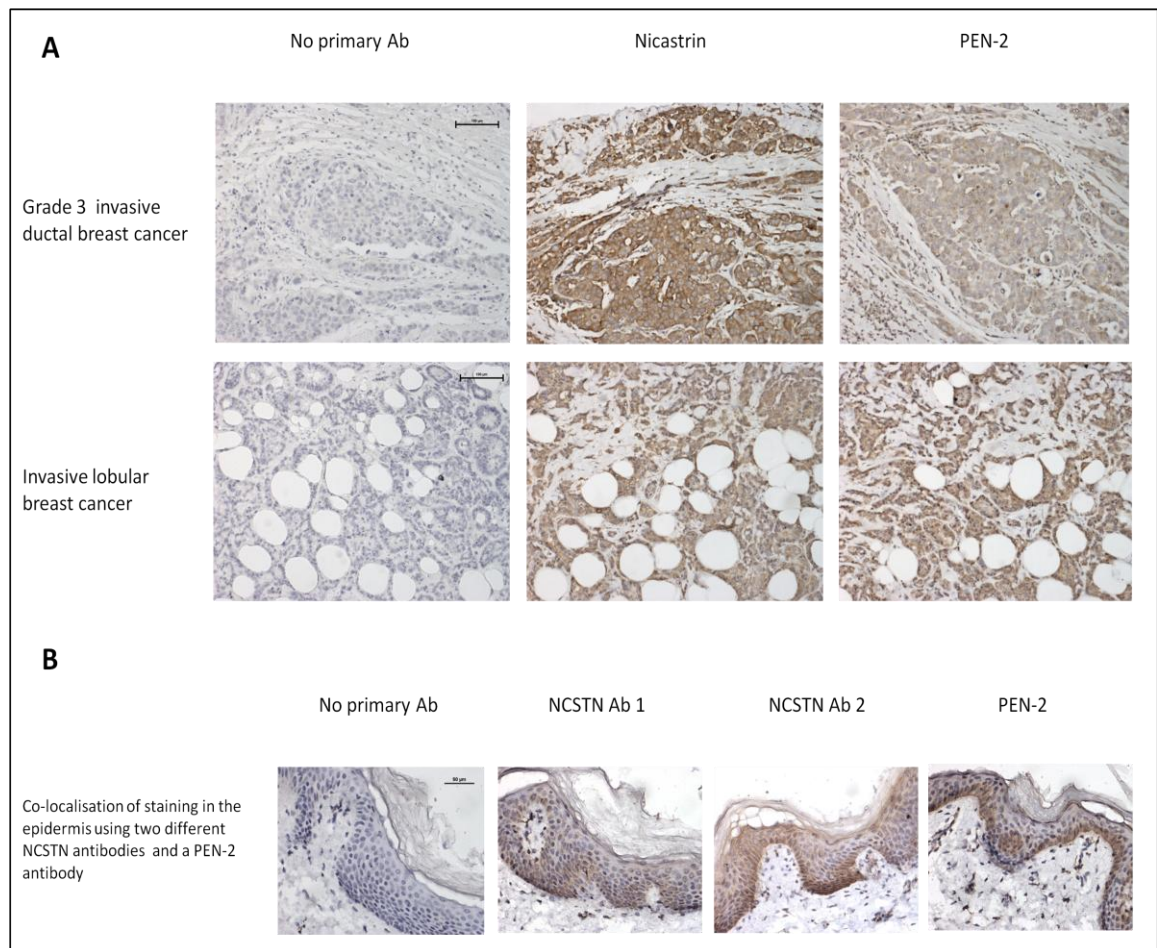
**Figure 4.35. PEN-2 expression in solubilised cell membrane fractions harvested from mutant and control fibroblasts.** Data represents the average protein band density measured across seven independent experiments. There was no significant difference in PEN-2 expression in *PSENEN* mutant fibroblasts versus controls (error bars = SE).

## **4.6 Immunohistochemical analysis of NCSTN and PEN-2 expression in the skin**

Gamma-secretase is integral to the cleavage of over 30 trans-membrane proteins and is widely expressed throughout the body. The cutaneous location of gamma-secretase complexes and their individual constituents has not been studied in detail. We sought to investigate this in normal skin, skin from mutation-positive patients and skin from mutation-negative HS patients. All immunohistochemistry was performed by Dr. Carl Hobbs, The Wolfson Centre for Age Related Diseases, King's College London.

### **4.6.1 Primary antibody optimisation**

The NCSTN and PEN-2 antibodies used for immunoblotting were used to perform immunohistochemistry on paraffin-embedded tissue. Antibody concentrations were chosen that produced clear and visible staining in skin with no apparent non-specific background signal (1:200 and 1:20 for nicastrin and PEN-2 respectively). Nicastrin expression is up regulated in tumour cell lines and tissue sections derived from invasive ductal breast carcinomas<sup>272</sup>. Tissue from a grade IV invasive and a lobular breast cancer (obtained from the Guy's Hospital Breast Cancer Tissue Bank) was therefore used as a positive control to assess antibody specificity. Nicastrin was abundantly expressed in both invasive and lobular breast cancer using the chosen nicastrin antibody, supporting antibody specificity (Figure 4.36a). As another integral subunit of the gamma-secretase complex, one might also expect PEN-2 to be over-expressed in these specimens. This was indeed the case, supporting PEN-2 antibody specificity (Figure 4.36a). Co-localisation of NCSTN and PEN-2 staining in healthy volunteer skin, presumably representing the location of gamma-secretase complexes, further supported antibody specificity (additionally replicated using a second NCSTN antibody, Figure 4.36b).

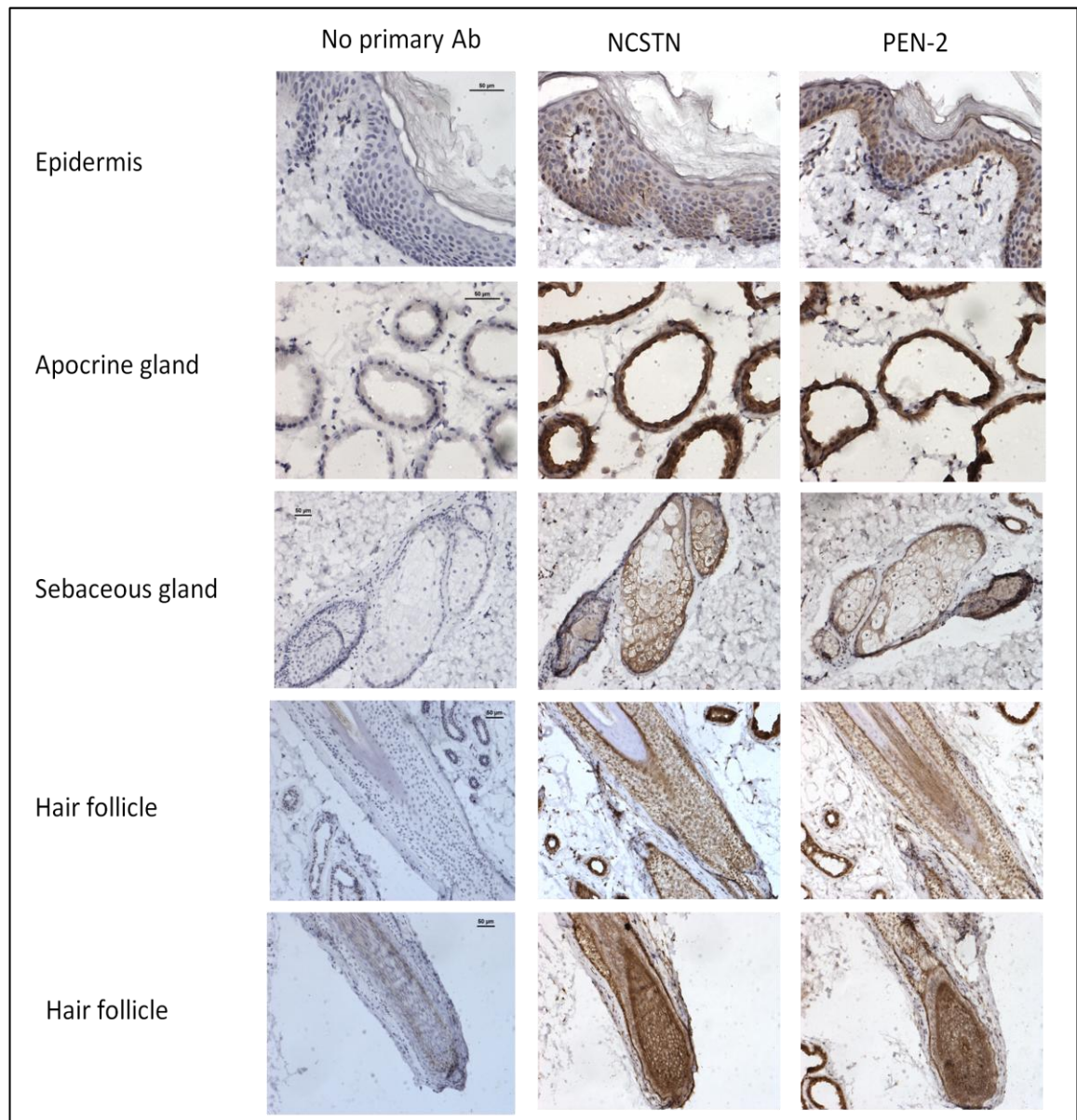


**Figure 4.36. Nicastrin and PEN-2 staining in human breast cancer and healthy skin.** No staining was seen in the no primary antibody run in breast cancer or skin. A/ Nicastrin and PEN-2 were highly expressed in grade 3 invasive ductal and lobular breast cancer (positive control). Scale bar = 100µm. B/ NCSTN (two separate primary antibodies) and PEN-2 expression co-localised in the epidermis, supporting antibody specificity. Scale bar = 50µm.

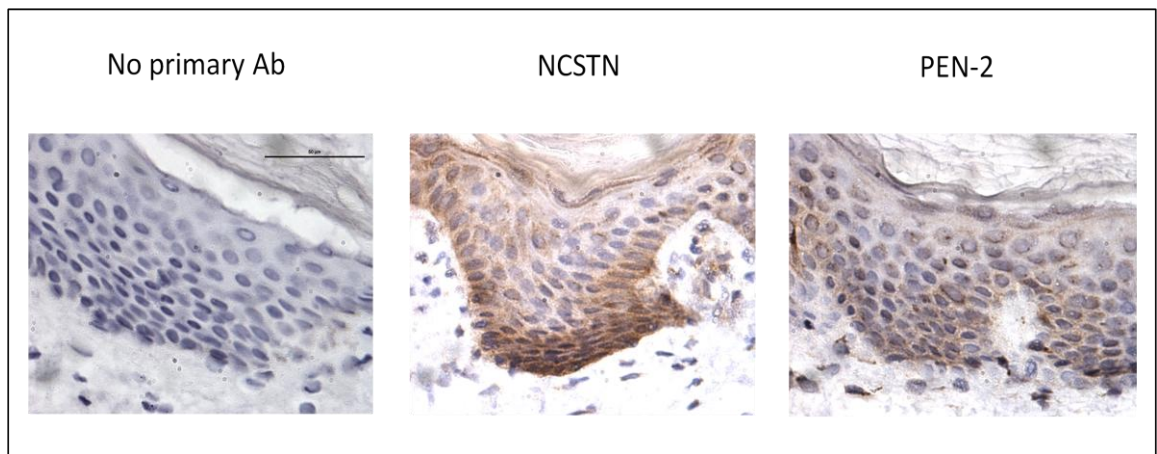
#### 4.6.2 Gamma-secretase is concentrated in the epidermis, hair follicle, apocrine gland and sebaceous gland in axillary skin harvested from healthy volunteers

Axillary skin was obtained from three healthy volunteer subjects, embedded in paraffin, sectioned and probed with the nicastrin and PEN-2 antibodies. Protein expression was most abundant in the epidermis, hair follicle, sebaceous gland and apocrine gland (Figure 4.37). Whilst Figure 4.37 represents only one of the volunteers, the pattern of staining was consistent across all three individuals. The extra-nuclear staining observed

in epidermal keratinocytes is consistent with the membrane-bound location of the gamma-secretase components (Figure 4.38).



**Figure 4.37. NCSTN and PEN-2 expression in healthy volunteer axillary skin.** NCSTN and PEN-2 staining co-localised in the epidermis, hair follicle, sebaceous gland and apocrine glands. Scale bar = 50µm.

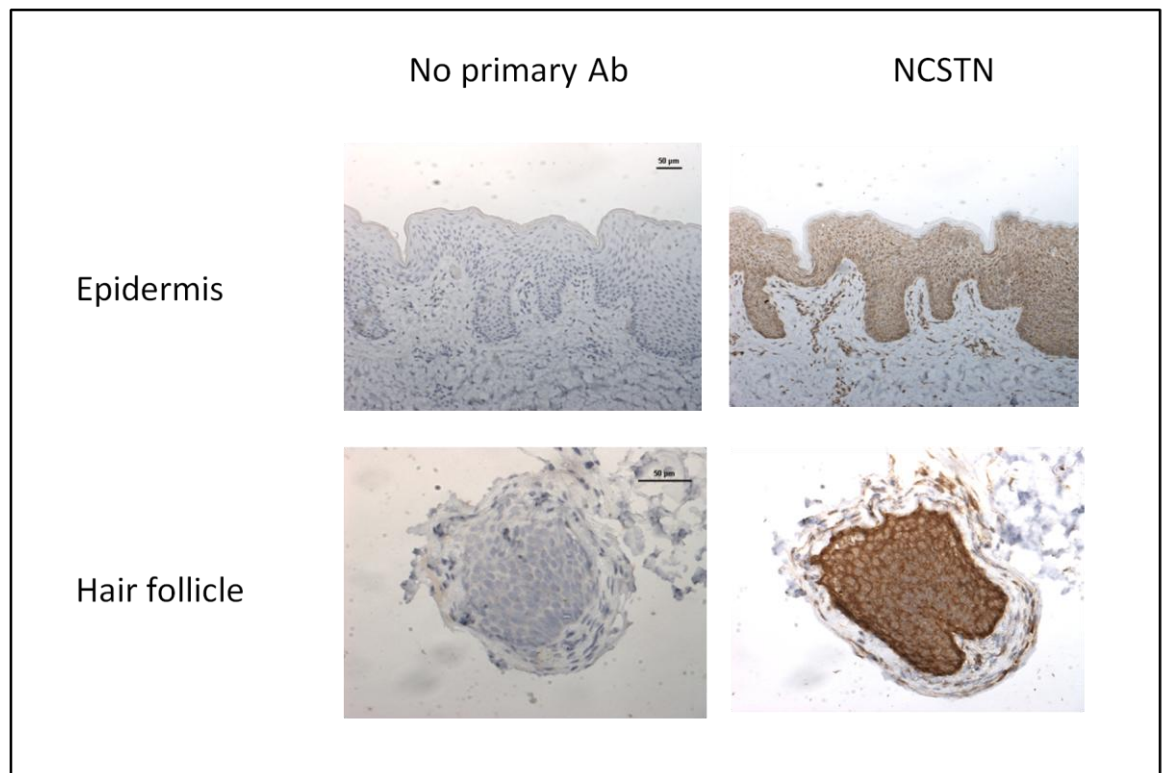


**Figure 4.38 Close up of NCSTN and PEN-2 staining in the epidermis.** Staining is extra-nuclear, consistent with the predominantly membrane-bound location of the gamma-secretase components in cells. Scale bar = 50µm.

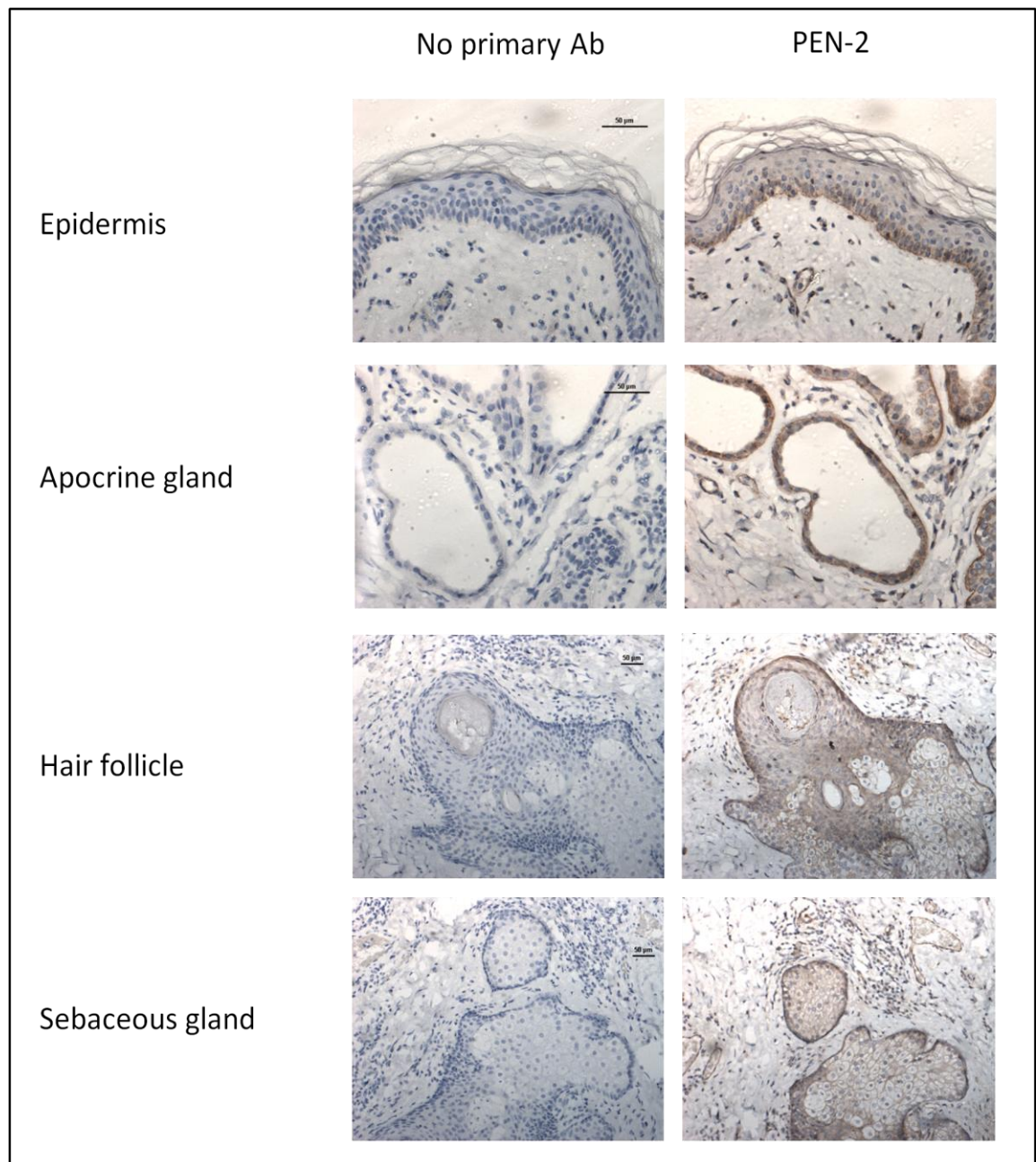
#### **4.6.3 NCSTN and PEN-2 are expressed in the epidermis, hair follicle, apocrine gland, sebaceous gland and inflammatory infiltrate in axillary skin harvested from patients with *NCSTN* c.1125+1 G>A and *PSENEN* c.66\_67insG mutations**

The expression pattern of nicastrin and PEN-2 in affected, site-matched, axillary skin from the two patients harbouring mutations in *NCSTN* and *PSENEN* (*NCSTN* c.1125+1 G>A, *PSENEN* c.66\_67insG) did not significantly differ from that observed in healthy volunteers, however additional staining was evident within the inflammatory infiltrate (Figure 4.39, Figure 4.40). Staining with antibodies directed against myeloperoxidase, CD3, L26 and CD68 suggested that neutrophils, T cells, B cells and macrophages were present within the inflammatory infiltrate.

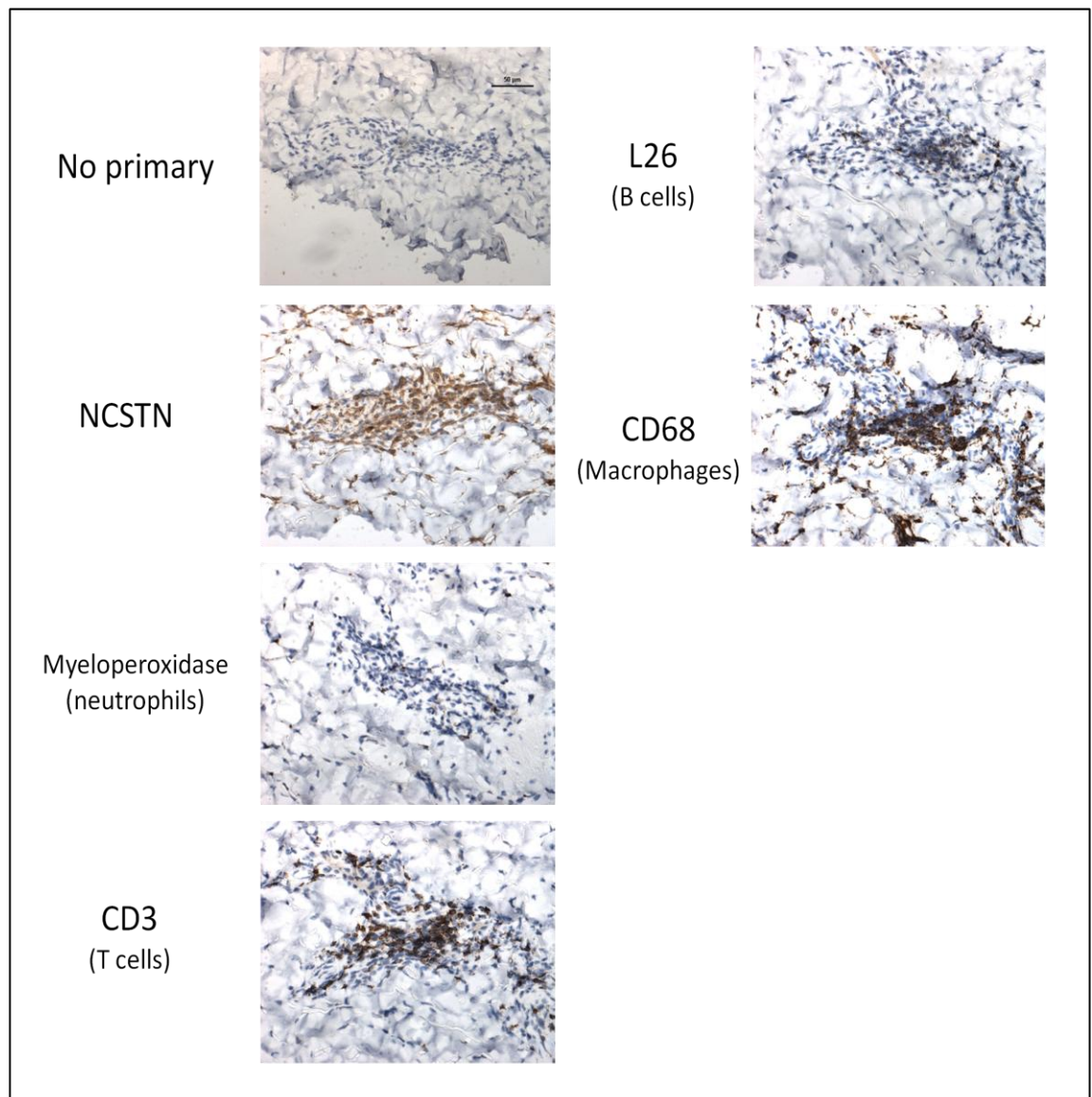




**Figure 4.39. Cutaneous NCSTN expression in affected axillary skin from the individual harbouring the *NCSTN* c.1125+1 G>A mutation.** Expression can be seen in the epidermis, hair follicle and inflammatory infiltrate. Scale bar = 50µm.

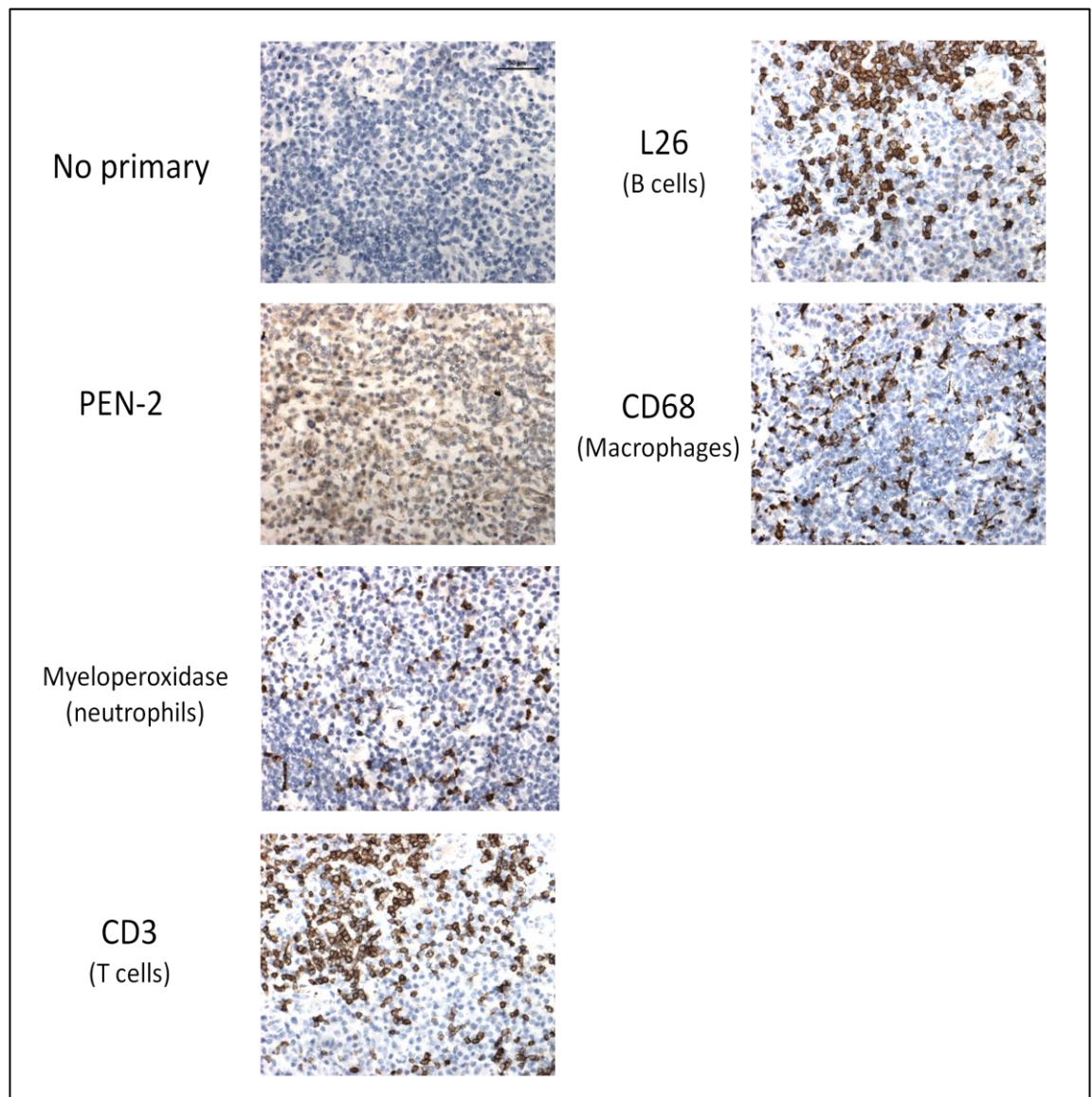


**Figure 4.40. Cutaneous PEN-2 staining in affected axillary skin from the individual harbouring the *PSENEN* c.66\_67insG mutation.** Expression can be seen in the epidermis, apocrine glands, hair follicle, sebaceous glands and the inflammatory infiltrate. Scale bar = 50μm.



**Figure 4.41. Immunohistochemical staining of the inflammatory infiltrate observed in axillary skin harvested from the patient harbouring the *NCSTN* c.1125+1 G>A mutation.** No staining was observed in the no primary run. Cells within the inflammatory infiltrate appeared to express NCSTN and comprised neutrophils, T cells, B cells and macrophages. Scale bar = 50μm.

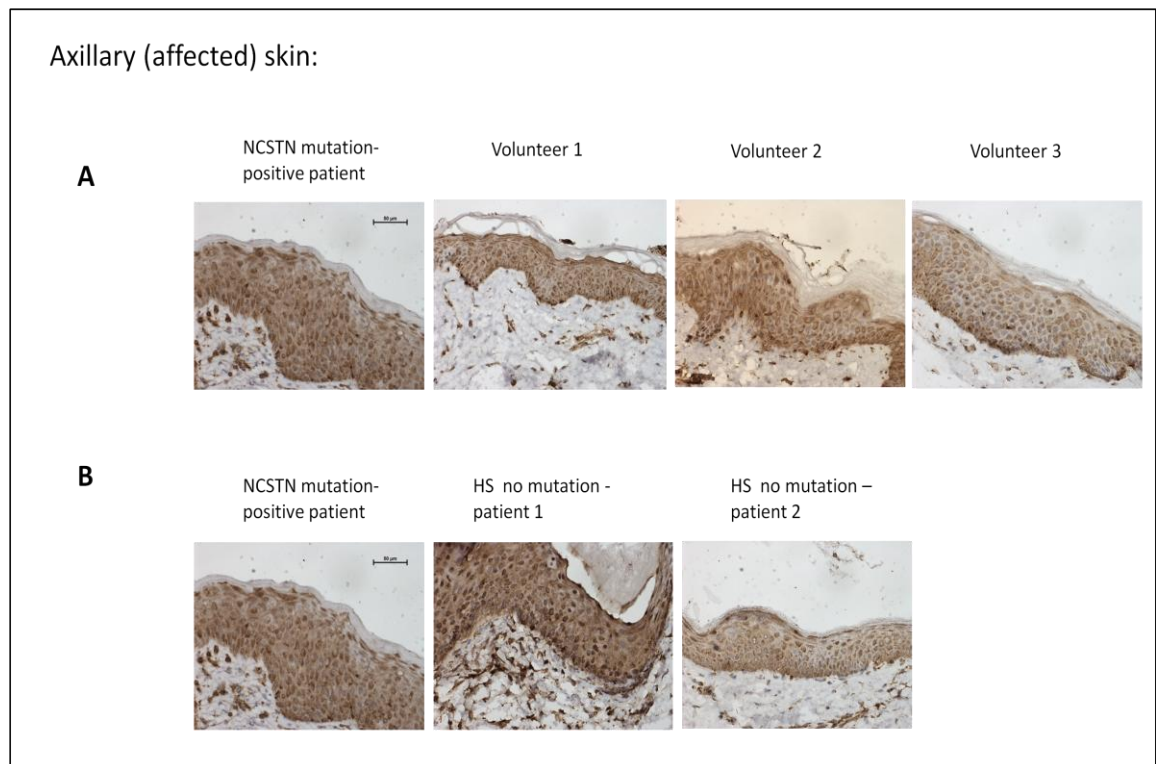




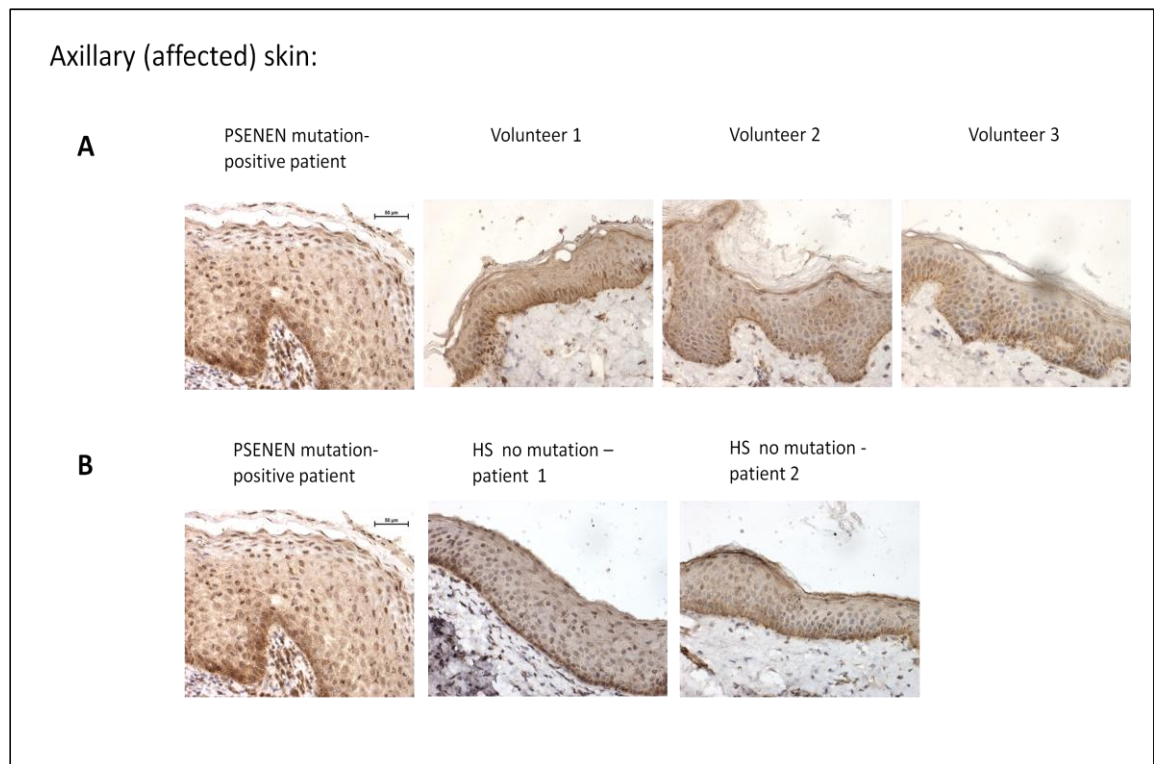
**Figure 4.42. Immunohistochemical staining of the inflammatory infiltrate observed in axillary skin harvested from the patient with the *PSENEN* 66\_67insG mutation.** No staining was observed in the no primary run. Cells within the inflammatory infiltrate appeared to express PEN-2 and comprised neutrophils, T cells, B cells and macrophages. Scale bar = 50µm.

#### **4.6.4 There was no discernible difference in epidermal Nicastrin and PEN-2 staining intensity in axillary skin harvested from patients with mutations versus healthy volunteers**

*NCSTN* c.1125+1 G>A and *PSENEN* c.66\_67insG were both associated with a reduced expression of their corresponding proteins (NCSTN and PEN-2) in primary human fibroblasts *in vitro* (sections 4.3.4, 4.4.4). It was therefore hypothesised that there would be a corresponding detectable reduction in protein expression in paraffin-embedded skin harvested from individuals harbouring the mutations. Axillary skin from the two mutation-positive patients and three healthy volunteers was analysed in tandem. No discernible difference in NCSTN or PEN-2 epidermal staining intensity was detected in axillary skin harvested from patients harbouring the mutations versus healthy volunteers (Figure 4.43a, Figure 4.44a). HS can dramatically distort the normal histological architecture of the skin, thus affected axillary skin from HS patients with wild-type *NCSTN*, *PSEN1* and *PSENEN* was studied as a second comparator group. Again there was no clear difference in NCSTN or PEN-2 expression in the epidermis of mutation-positive patients versus wild type HS subjects (Figure 4.43b, Figure 4.44b). The effect of local inflammation on gamma-secretase expression is unknown. To account for this, para-axillary skin (no clinically apparent inflammation) from the patient with the *PSENEN* mutation was compared against matched para-axillary skin from three healthy volunteers. No clear difference in PEN-2 expression was observed (Figure 4.45). Para-axillary skin was not available from the patient harbouring the *NCSTN* mutation.

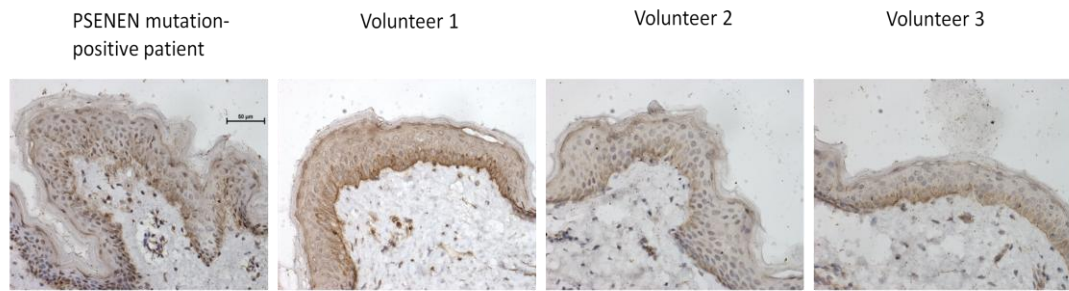


**Figure 4.43. A comparison of NCSTN epidermal staining in axillary skin harvested from the patient with an *NCSTN* c.1125+1 G>A mutation and control subjects.** A) Affected axillary skin from the patient harbouring the *NCSTN* c.1125+1 G>A mutation is compared against site-matched axillary skin harvested from healthy volunteer subjects. There was no discernible difference in protein expression in the mutation-positive patient versus controls. Scale bar = 50μm B) Affected axillary skin from the patient harbouring the *NCSTN* c.1125+1 G>A mutation is compared against affected axillary skin harvested from two HS patients wild-type for the *NCSTN*, *PSENEN* and *PSEN1* genes. Again, there was no discernible difference in NCSTN expression between individuals. Scale bar = 50μm.



**Figure 4.44. A comparison of PEN-2 epidermal staining in axillary skin harvested from the patient harbouring the *PSENEN* c.66\_67insG mutation and healthy volunteers.** A) Affected axillary skin from the patient harbouring *PSENEN* c.66\_67insG is compared against site-matched axillary skin harvested from healthy volunteer subjects. There was no discernible difference in protein expression in the mutation-positive patient and healthy volunteers. Scale bar = 50µm. B) Affected axillary skin from the patient harbouring *PSENEN* c.66\_67insG is compared against affected axillary skin harvested from two HS patients wild-type for the *NCSTN*, *PSENEN* and *PSEN1* genes. Again, there was no significant difference in PEN-2 expression between individuals. Scale bar = 50µm.

Para-axillary (unaffected) skin:



**Figure 4.45. A comparison of PEN-2 epidermal staining in para-axillary skin harvested from the individual harbouring *PSENEN* c.66\_67insG and healthy volunteers.** There was no discernible difference in PEN-2 expression in the epidermis of the individual harbouring the mutation versus healthy volunteers. Scale bar = 50µm.

## 4.7 Discussion

This aim of this chapter was to functionally characterise the two gamma-secretase mutations identified in familial cases of HS in Chapter 3 (*NCSTN* c.1125+1 G>A, *PSENEN* c.66\_67insG). The two mutations confer very different predicted effects on their respective transcripts, one causing aberrant splicing, the other resulting in a frameshift, however they both appear to share a common downstream pathogenic mechanism. Both mutations result in a reduced transcript abundance and a corresponding reduction in protein expression, highlighting haploinsufficiency as the likely common pathogenic mechanism. This is consistent with functional data generated from some other globally reported gamma-secretase mutations. Peripheral lymphocyte mRNA expression was assessed in Chinese, Japanese and French individuals with gamma-secretase mutations. The following nonsense, frameshift and splice site mutations were associated with reduced transcript abundance, *NCSTN* p.R117X, *NCSTN* c.582+1delG, *NCSTN* p.Q163SfsX39, *PSEN1* p.P242LfsX11<sup>135; 139; 141</sup>. In contrast, some splice site, frameshift and even nonsense mutations did not affect transcript abundance (*NCSTN* p.R434X, *NCSTN* p.S590AfsX3, *NCSTN* p.E584DfsX44 and *NCSTN* c.1551+1 G>A). One would hypothesise that NMD underlies the apparent transcript reduction in nonsense, frameshift and splice site mutations which result in premature stop codons. In the case of the two mutations reported here, and potentially *NCSTN* c.582+1delG reported in a Japanese patient (exact splicing defect not characterised), there is no evidence to suggest that premature stop codons are formed, thus making NMD an unlikely mechanism of RNA decay. *NCSTN* c.1125+1 G>A is predicted to alter the secondary RNA structure which may render it unstable and subject to decay. Alternatively the altered structure may inhibit efficient ribosomal processing and potentially induce no-go decay<sup>271</sup>. *PSENEN* c.66\_67insG is not predicted to significantly alter the secondary RNA structure but the altered reading frame does

eliminate the native stop codon, potentially resulting in non-stop decay. This is however a translation-associated mechanism of RNA decay and as such, one might have expected cycloheximide to have affected RNA abundance, which was not the case. The exact mechanism of RNA decay in these cases requires further investigation<sup>270</sup>. Both mutations potentially provide a platform upon which to further our current understanding of RNA surveillance and decay mechanisms.

#### **4.7.1 Both mutations appear to result in haploinsufficiency but the respective protein expression at the cell membrane remains unaffected**

Both mutations led to a reduction in respective protein expression in total cell fractions but this was not replicated in solubilised cell membrane preparations. This might suggest that, despite a reduction in the total amount of protein available, there are sufficient levels to support normal protein expression at the cell membrane in the *in vitro* conditions studied. Gamma-secretase gene transcription and protein expression is complex and tightly regulated. It has been demonstrated that *NCSTN*, *PSENEN* and *PSEN1* mRNA abundance does not closely correlate with protein expression in the liver and heart. High mRNA levels are associated with very limited expression of their respective proteins and the authors proposed that this may be due to additional post-transcriptional mechanisms<sup>160</sup>. Even following effective translation, fewer than 5% of gamma-secretase complexes that are assembled in the ER travel to the Golgi and are distributed to cell membranes. Proteins that are not incorporated into complexes are rapidly degraded<sup>189</sup>. This might suggest that even in a situation of heterozygous haploinsufficiency, the 50% wild type protein that is still present may be sufficient to support normal enzyme complex formation.

It should be noted that these conclusions are drawn from immunoblotting experiments which are only semi-quantitative. Furthermore, immunoblots of solubilised cell membrane protein samples were analysed in the absence of a reliable loading control.

This was partially circumnavigated by generating consistent sample concentrations using BCA and repeating experiments seven times to reduce any human error (eg. pipetting). This does however rely on the accuracy of BCA in measuring solubilised cell membrane preparation concentrations in which many proteins, including NCSTN and PEN-2, are incorporated in complexes. The efficiency of cell membrane protein solubilisation, particularly the solubilisation of gamma-secretase complexes, is another important variable in this study.

#### **4.7.2 These mutations are not associated with a significant difference in gamma-secretase complex number, structure or activity**

Mutations have now been reported in three of the four gamma-secretase enzyme complex components (*NCSTN*, *PSENEN* and *PSEN1*) in HS. This implies that a consequential effect on gamma-secretase enzyme function may be the common pathogenic mechanism underlying this form of HS. Enzyme activity assays were therefore performed to assess gamma-secretase activity in primary dermal fibroblasts *in vitro*. There was no significant difference in PSEN1-CTF expression, representative of mature gamma-secretase complex number, in mutant fibroblasts versus controls. Complex maturation, as measured by the PSEN1-CTF/ PEN-2 ratio also appeared unaffected. It should be noted however that whilst this is a common and accepted way to measure complex maturity, one of the patients harbouring a mutation had a mutation in *PSENEN*. Any consequent lack of PEN-2 protein may therefore directly influence the PSEN1-CTF/ PEN-2 ratio. This is further confounded by the fact that any reduction in PEN-2 may also prevent the cleavage and maturation of PSEN-1 (thus reducing the total amount of PSEN1-CTF present) which could also alter, or even revert, the final PSEN1-CTF/ PEN-2 ratio. This result should therefore be interpreted with caution in the individual harbouring a mutation in *PSENEN*. Contrary to the hypothesis that enzyme activity may be reduced in mutant fibroblasts, neat endopeptidase activity per mg of



solubilised cell membrane protein was unaltered in *NCSTN* mutant fibroblasts and appeared slightly increased in *PSENEN* mutant fibroblasts. Endopeptidase activity per enzyme complex did not appear to differ in *PSENEN* mutant fibroblasts versus controls however. There was no significant difference in carboxypeptidase activity, as measured by A $\beta$ 42: A $\beta$ 40 ratios, but it was noticeable that the assays run with samples harvested from the individual with a *PSENEN* mutation generated lower ratios than all three controls, potentially indicative of marginally increased carboxypeptidase activity. Overall, there was no highly significant difference in complex number, maturity or activity in mutant fibroblasts versus controls. The slightly increased endopeptidase activity per mg solubilised cell membrane protein and the slightly lower A $\beta$ 42: A $\beta$ 40 ratios observed in the individual harbouring a *PSENEN* mutation require further investigation and would need to be validated against a larger number of controls. Any such effect would provide an intriguing contrast to the enzymatic disruption arising from *PSEN1* and *PSEN2* mutations in FAD.

These enzyme activity assay results are associated with all of the limitations of *in vitro* cell experimentation. Cells in a tissue culture environment are not truly representative of their physiological state. Furthermore, large numbers of cells were required for the enzyme activity assays and recurrent cell passaging (up to 10 passages) may have significantly affected cell behaviour and function. Despite growing similar numbers of cells there was marked variability in the concentration of solubilised cell membrane protein generated from each individual. This may be due to variation in the efficiency of the protein extraction method or due to subtle differences in general cell culture techniques (eg. cell confluency, passaging and general culture conditions). Enzyme activity assays are sensitive and results can vary depending upon a multitude of factors including the quality of reagents used, pipetting error and variable assay conditions. To

reduce this variability, reagents from the same batch were used for all experiments and all assays were repeated seven times.

#### **4.7.3 Cutaneous expression of the gamma-secretase components correlates with the histopathological changes observed in HS, supporting a role for gamma-secretase mutations in disease pathogenesis**

Immunohistochemistry revealed that NCSTN and PEN-2 co-localise in healthy axillary skin, as would be expected if the staining is representative of gamma-secretase complex location, and are predominantly expressed in the epidermis, hair follicle, sebaceous glands and apocrine glands. The widespread expression was not unexpected given the integral role of gamma-secretase in maintaining essential cell signalling pathways such as Notch signalling. It is noteworthy however that the areas of prominent staining correlate with the characteristic histo-pathological changes observed in HS. These comprise follicular plugging, follicular dilatation, peri-follicular inflammation, apocrine gland stasis and peri-glandular inflammation, inter-follicular epidermal hyperplasia, sebaceous gland involution and potentially thinning of the BMZ at the sebo-follicular junction<sup>6; 7; 27; 28; 80</sup>. The distribution of staining therefore potentially supports a role for gamma-secretase in disease pathogenesis. The only significant difference observed between mutation-positive individuals and healthy volunteers was the additional expression of the gamma-secretase components in the mixed inflammatory infiltrate. There was no evidence on IHC to support a quantitative difference in NCSTN or PEN-2 expression in individuals harbouring mutations versus healthy volunteers. This should be interpreted with caution however as there is always a degree of variability in immunohistochemistry (IHC) results due to the method of staining and the nature of the sections studied. For example, in contrast to sections from healthy volunteers, HS affected skin may comprise inflammatory infiltrates and will exhibit many of the typical histopathological changes mentioned above. Furthermore, IHC is not a quantitative

technique, thus protein expression cannot be interpreted with any certainty from this study.

The above functional studies used primary dermal fibroblasts as a source of RNA, protein and gamma-secretase enzyme because they express the gamma-secretase components and are stable and sustainable in cell culture, a factor that was of significant importance given the number of cells required to perform the enzyme activity assays. At the time of study, it was unclear which cutaneous cell type was most relevant in HS. It was subsequently noteworthy that immunohistochemistry revealed significantly enhanced expression of gamma-secretase components in keratinocytes compared with dermal fibroblasts, potentially highlighting keratinocytes as another cell type worthy of future study.

#### **4.7.4 Conclusions**

The *NCSTN* and *PSENEN* mutations studied here appear to result in haploinsufficiency and a corresponding reduction in protein expression in total cell fractions. They would not however appear to significantly affect protein expression and cause effective haploinsufficiency at the level of the cell membrane. There would appear to be no detrimental effect on complex maturity, number and activity at membrane level however those results only represent the complexes that have reached that site and have been effectively solubilised. If haploinsufficiency is indeed the primary pathogenic mechanism then attention may need to be focussed on gamma-secretase complex formation and trafficking at earlier stages of cell cycling. Any reduction in the levels of wild type NCSTN or PEN-2 in the ER may affect the quantity of the other gamma-secretase constituents. As discussed in the introduction, the gamma-secretase components rely on one another for stability. Therefore through an independent or compound effect on the other components, a reduction in NCSTN or PEN-2 expression may detrimentally affect complex formation. This may result in altered, unstable and

dysfunctional complexes. As previously mentioned, only around 5% complexes assembled in the ER are transported to the golgi for maturation and are exported to the membranes<sup>189</sup>. It is therefore important to ascertain whether there is any alteration in the number and nature of complexes exported from the ER. There is no evidence to support the presence of altered complexes within the cell membrane however that is not to say that they are not present and potentially functional at that site or in sub-membranous locations. They may be present within the cell membrane but degraded by the solubilisation process employed in this study, thus their enzymatic effects may be underestimated. Solubilised cell membrane protein extracted from mouse cells has previously been shown to contain bound and unbound gamma-secretase complex components suggesting that even stable WT complexes may be disrupted during the solubilisation process<sup>235</sup>.

An intriguing feature of HS is the very characteristic cutaneous distribution of disease. Working on the basis that there may be less protein present at the level of the ER one might imagine that in certain circumstances or environments there may be a higher turnover of complexes required, thus putting more strain on the production and maturity of complexes. It is under strained conditions that a 50% reduction in protein would become more significant and potentially functionally detrimental. These cases represent interesting, and potentially very revealing, models of gamma-secretase function as the vast majority of *in vitro* and *in vivo* experimentation has involved studying the consequences of complete knockout or knockdown of various complex components. The subtleties of partial reductions in protein expression need to be investigated. It has previously been demonstrated in breast cancer cells that >90% NCSTN knockdown only results in a 50% reduction in gamma-secretase activity, demonstrating the resilient capability of cells to produce and maintain enzyme function<sup>272</sup>. The anatomical location of HS might suggest that intrinsic and extrinsic factors such as temperature, sweat and

friction may potentially impact upon complex turnover or function. Temperature, pH and salt concentrations have previously been shown to affect endopeptidase and carboxypeptidase enzyme activity *in vitro*<sup>266</sup>. It may therefore be advantageous to perform enzyme activity assays on mutant and control cells under different forms of stress to look for any subtle alteration in enzyme activity. After all, if mutation-positive patients had a significant inherent gamma-secretase dysfunction one might assume that it would confer more damaging effects than those observed in HS.

As mentioned previously, not all globally reported gamma-secretase mutations appear to result in haploinsufficiency. That is not to say that protein levels are unaffected however, as they were not studied. On the basis of this work, and assuming that there is a common pathogenic mechanism underpinning gamma-secretase gene related HS, one might imagine that those mutations result in loss of function. It is noteworthy that in the cases studied here there was some, however minimal, mutant transcript expressed. There was no evidence to support the production of mutant protein on immunoblotting but the molecular weight of the predicted mutant proteins would not significantly differ from the wild-type proteins and would not be easily visible. One cannot therefore exclude the possibility that mutant protein may be conferring some effect in disease pathogenesis in these cases and this requires further investigation.

An important limitation of this study was that the majority of tissue samples were harvested from affected axillary skin in mutation-positive patients and site-matched unaffected axillary skin in healthy volunteers. As previously discussed, the effect of active disease and the associated inflammation (present in affected skin) on RNA, protein and enzyme activity is unknown. Whilst *in vitro* cell culture and the accompanying cell passages are likely to negate any such effects this does remain an unknown variable. It is clearly important to assess samples taken from affected skin but future studies would potentially benefit from analysing additional samples harvested

from para-axillary unaffected skin which would provide a more like for like resting state comparison. Another limitation of this study was that whilst all healthy volunteer subjects (used as controls) reported no personal or family history of HS or associated conditions they were not formerly screened for gamma-secretase gene mutations to ensure that they were truly wild type controls. In support of this, novel gamma-secretase mutations are extremely rare and have not been observed in over 600 in-house controls who have been subjected to exome sequencing. It was therefore very unlikely that any of these three individuals harboured any such variants. Furthermore, all three behaved in a consistent and concordant manner throughout these studies.

This chapter sought to functionally characterise two gamma-secretase mutations identified in multiplex kindreds with HS. These data represent an important step in starting to understand the cellular mechanisms underlying disease formation in these patients and provide a platform upon which to design further studies. The identification and investigation of a more diverse array of mutations may further facilitate that process. Ongoing functional work is clearly required to make this genetic discovery truly translational. To understand the importance of these functional studies in the context of HS as a whole it was important to ascertain the prevalence of gamma-secretase mutations in the general disease population. This is addressed in Chapter 5.

## **Chapter 5: Results: Mutations in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1* underlie a minority of cases of hidradenitis suppurativa**

### **5.1 Introduction**

Gamma-secretase mutations had been reported in 11 multiplex kindreds worldwide (Chinese, British and Japanese kindreds)<sup>135-138</sup> yet, as exemplified by our early familial study (Chapter 3), they did not appear to account for all familial cases of HS<sup>139</sup>. Ascertaining the prevalence of *NCSTN*, *PSENEN* and *PSEN1* mutations was therefore of paramount importance both in terms of future genetic study and the translational relevance of these findings.

#### **5.1.1 Aims and objectives**

The aim of this work was to determine the prevalence of mutations in each of these genes amongst a large cohort of subjects with HS, representative of the general disease population.

## **5.2 Patient cohort**

Forty eight individuals with HS were sequentially consented and recruited from the tertiary referral clinic at St. John's Institute of Dermatology. The cohort was of mixed ethnicity, 65% were female, the average body mass index was 31.1 (clinically obese), 69% reported a smoking history and the average Sartorius score was 49.9 (see Table 5.1). 20 (42%) patients reported a family history of HS, 19 consistent with autosomal dominant inheritance.

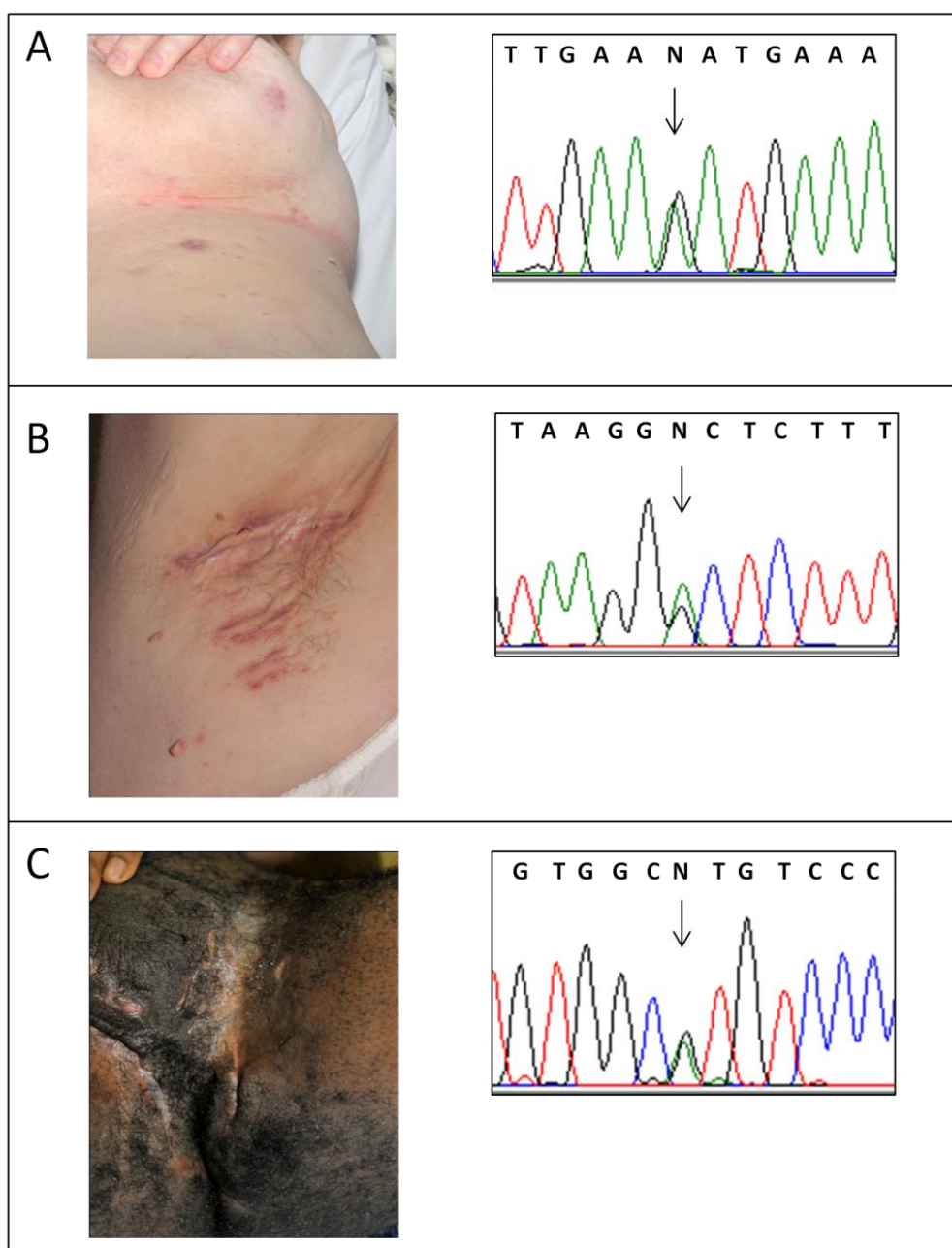


**Table 5.1. Population and clinical data on the 48 HS patients recruited to the study**

<b>Sex:</b>	
<b>Male</b>	17/ 35%
<b>Female</b>	31/ 65%
<b>Ethnicity:</b>	
<b>Caucasian</b>	28/ 58%
<b>Afro-Caribbean</b>	15/ 31%
<b>Asian</b>	5 / 11%
<b>Body mass index</b>	31.1 (19 – 58)
<b>Smoking history</b>	33/ 69% (48% started smoking prior to disease development)
<b>Age of onset</b>	21.4 (5 – 48)
<b>Dermatology life quality index</b>	14.87 (1 – 30)
<b>Sartorius severity score</b>	49.9 (11-156)
<b>Most common sites affected:</b>	
<b>Groin</b>	47/ 98%
<b>Axillae</b>	45/ 94%
<b>Buttocks</b>	38/ 79%
<b>Chest</b>	16/ 34%
<b>Posterior ears</b>	9/ 19%
<b>History of pre-menstrual flares</b>	14 of 31 women (45% women)
<b>Family history of HS</b>	20/ 42%
<b>History of acne</b>	11/ 23%

### **5.3 Three of forty eight individuals harboured novel variants in the gamma-secretase genes *NCSTN*, *PSEN1* and *PSENEN***

All coding regions and associated splice sites of *NCSTN*, *PSEN1* and *PSENEN* were amplified by PCR, using exon flanking intronic primers, and Sanger sequenced in all 48 individuals. The whole cohort was assessed for large scale deletions and duplications in *NCSTN*, *PSENEN* and *PSEN1* by multiplex ligation-dependent probe amplification (MLPA). Three subjects harboured novel heterozygous DNA changes in *NCSTN* comprising a missense substitution (*NCSTN* c.553G>A, p.Asp185Asn) and two single base substitutions located within ten base pairs of donor splice junctions (*NCSTN* c.996+7G>A and *NCSTN* c.1101+10A>G) (Figure 5.1). No novel or rare variants (population frequency < 1%, defined as a population frequency of <1% in both the 1000 Genomes Project database<sup>260</sup> and NHLBI Exome Sequencing Project EVS (<http://evs.gs.washington.edu/EVS/>)) were detected in coding regions or within 10 base pairs of splice boundaries in the remaining 45 individuals. No CNV was detected in *PSEN1*, *PSENEN* or *NCSTN* via MLPA.



**Figure 5.1. Three individuals with novel variants in *NCSTN*.** A) Patient harbouring *NCSTN* c.553 G>A. Inflammatory nodules, abscesses and scarring in the sub-mammary region. Arrow indicating the *NCSTN* c.553 G>A variant, a novel heterozygous substitution in exon 5 of *NCSTN*. B) Patient harbouring *NCSTN* c.996+7 G>A. Inflammatory nodules and extensive scarring in the axilla. Arrow indicating the *NCSTN* c.996+7 G>A variant, a heterozygous substitution of the seventh base of the donor splice site within intron 8 of *NCSTN*. C) Patient harbouring *NCSTN* c.1101+10. Inflammatory nodules, sinus tracts and scarring in the groin. Arrow indicating the *NCSTN* c.1101+10 A>G variant, a heterozygous substitution of the tenth base of the donor splice site within intron 9 of *NCSTN*. Photos courtesy of Guy's and St. Thomas' NHS Foundation Trust, all individuals consented to publication.

## 5.4 Identification of a novel missense variant in exon 5 of *NCSTN* (c.553 G>A, p.Asp185Asn)

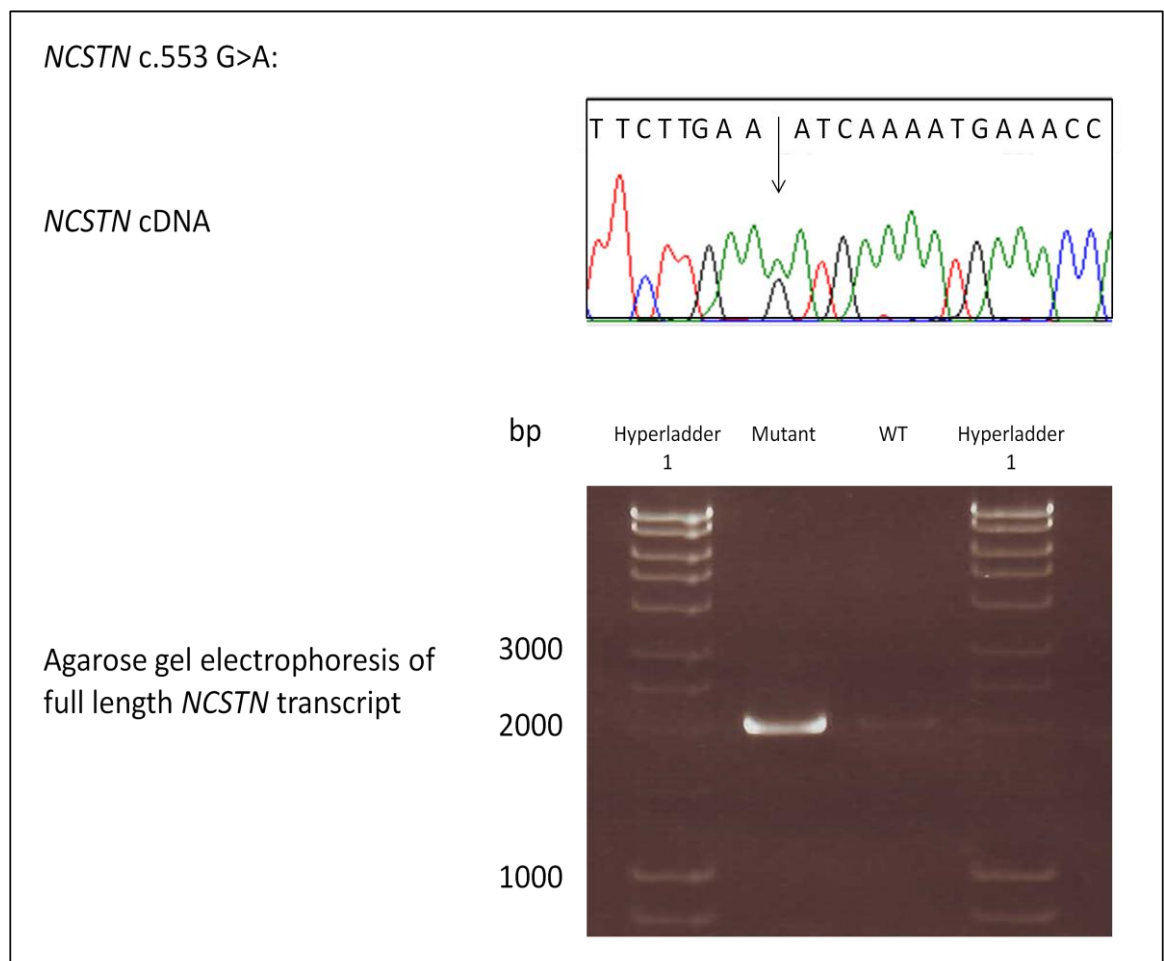
A novel heterozygous missense variant in exon 5 of *NCSTN* (c.553G>A, p.Asp185Asn) was identified in a 45 year old female (Figure 5.1a) with no family history of HS. This is predicted to result in the substitution of an acidic, polar and negatively charged aspartic acid residue with a polar and neutral asparagine residue. This residue is conserved down to rabbit, in which, notably, the aspartic acid residue is substituted with an asparagine residue (Figure 5.2). The SIFT and PolyPhen-2 algorithms for predicting the functional consequences of non-synonymous variants indicate that this variant is unlikely to be of functional significance (SIFT score 0.2, PolyPhen score 0.008)<sup>276; 277</sup>. The variant was not observed in 200 control subjects of European origin.

Human	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	Q	D	H	N	L	-
Mole rat	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	Q	D	H	N	L	-
Guinae pig	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	I	I	K	-	Q	C	Y	Q	D	H	N	L	-
Giant panda	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Pig	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	D	E	T	-	-	N	I	I	K	-	Q	C	Y	R	D	H	N	L	-
Horse	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Dog	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	H	D	H	N	L	-
Little brown rat	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Mouse	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	Q	D	H	N	L	-
Rabbit	L	A	Y	E	D	F	S	F	P	I	F	L	I	K	N	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	H	D	H	N	L	-
Cow	L	A	Y	D	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	Q	D	H	N	L	-
Hamster	L	A	Y	E	D	F	S	F	P	I	F	L	I	K	N	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	H	D	H	N	L	-
Opossum	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Tasmanian devil	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Turkey	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	K	V	I	K	-	Q	C	Y	Q	D	H	N	L	-
Chicken	L	S	Y	E	D	F	D	F	P	I	F	L	L	E	D	A	N	E	T	-	-	Q	V	I	K	-	Q	C	Y	Q	D	H	N	V	-
Finch	L	S	Y	E	D	F	D	F	P	I	F	L	L	E	D	A	N	E	T	-	-	Q	V	I	K	-	Q	C	Y	Q	D	H	N	V	-
Frog	L	S	Y	E	D	F	D	F	P	I	F	L	L	E	D	A	N	E	T	-	-	Q	V	I	K	-	Q	C	Y	Q	D	H	N	V	-
Elephant	L	S	Y	E	D	F	D	F	P	I	F	L	L	E	D	A	N	E	T	-	-	Q	V	I	K	-	Q	C	Y	Q	D	H	N	V	-
Zebrafish	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Tilapia nilotica	L	A	Y	E	D	F	S	F	P	I	F	L	L	E	D	E	N	E	T	-	-	N	V	I	K	-	Q	C	Y	R	D	H	N	L	-
Stickleback	L	S	Y	E	D	F	T	F	P	V	F	A	L	K	D	E	N	Q	T	-	-	Q	V	I	R	-	K	C	Y	E	D	H	N	L	-
Pufferfish	L	S	Y	E	D	F	T	F	P	V	F	A	L	K	D	E	N	Q	T	-	-	Q	V	I	R	-	K	C	Y	E	D	H	N	L	-

**Figure 5.2. Conservation of the aspartic acid residue (D185) affected by the *NCSTN* c.553 G>A variant.** The amino acid is not fully conserved in mammals. The aspartic acid residue is replaced by an asparagine residue in rabbit and hamster.

This lady reported painful nodules, cysts, abscesses and resultant scarring in her axillae, sub-mammary region, abdomen, groin and buttocks. She had a BMI of 38 (clinically obese) and was an active smoker. Her disease had failed to respond to multiple antibiotics including lymecycline, rifampicin and clindamycin, cyproterone acetate, acitretin and dapsone. Surgical excision of her axillae was effective.

Three missense mutations have since been reported in *NCSTN* in HS (c.223G>A, c.632C>G, c.647A>C<sup>138; 140</sup>). One further non-synonymous substitution (c.1768A>G) was found to result in aberrant splicing by creating a cryptic splice donor site<sup>139</sup>. Given this, we hypothesised that the c.553 G>A substitution reported here may also affect splicing. Computational splice site analysis did not predict that this variant would affect splicing (all splice site predictions performed using the Automated Splice Site and Exon Definition Analysis (ASSEDA) server<sup>278</sup>). RNA was collected from this patient and reverse-transcribed to cDNA. The full length *NCSTN* transcript was amplified by PCR and studied alongside that of a control subject by agarose electrophoresis. Only one band was visible in both cases (see Figure 5.3). The full length *NCSTN* transcript was Sanger sequenced and the sequence revealed the non-synonymous c.553G>A variant but no other abnormalities. These results do not imply that this missense variant affects splicing.



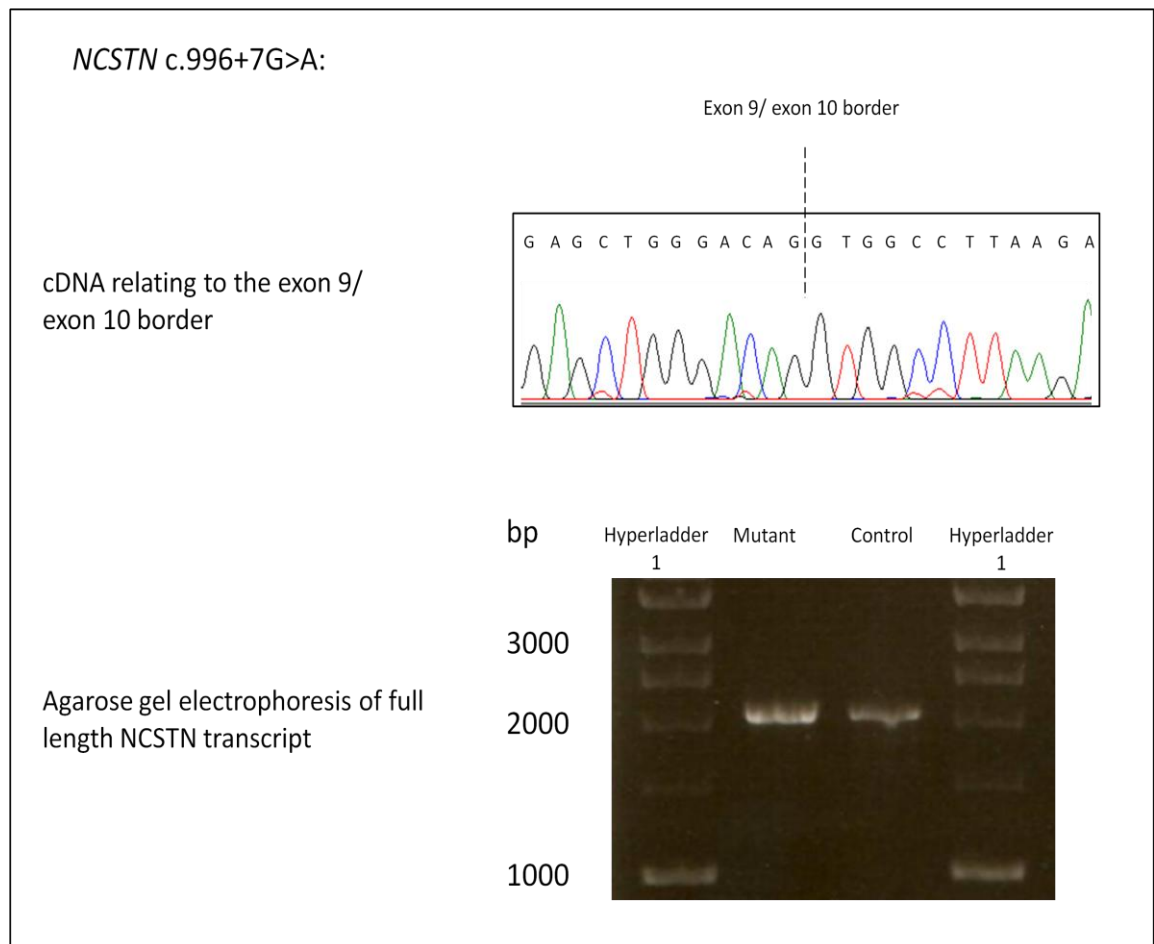
**Figure 5.3. Sequence and agarose electrophoresis of the full length *NCSTN* transcript in the individual with an *NCSTN* c.553 G>A variant.** The substitution is present in the full length *NCSTN* transcript but the cDNA sequence was otherwise normal. No extra bands were detected on agarose electrophoresis of the full length *NCSTN* transcript to imply that the variant affects splicing.

Overall, whilst this substitution is predicted to result in a change in acidity and charge of the corresponding amino acid residue it is not evolutionarily conserved, the exact same amino acid change is observed in rabbit, it is not predicted to be functionally damaging and it does not appear to affect splicing. The variant is therefore unlikely to be pathogenic.

## **5.5 Identification of a novel heterozygous substitution in the exon 8/ intron 8 splice site of *NCSTN* (c.996+7G>A)**

A novel heterozygous substitution of the conserved seventh base of intron 8 of *NCSTN* (c.996+7G>A) was identified in a 38 year old female (Figure 5.1b). She did not report a family history of the condition however it is noteworthy that both of her parents died when she was a child and two of her five siblings have each had one pilonidal abscess. The substitution was not observed in 200 controls of European origin. Computational splice site analysis predicts that this substitution may confer a detrimental effect on splicing (ASSED server<sup>278</sup>). RNA was extracted from whole blood lymphoblasts and converted to cDNA as previously described. No mutant transcript was evident on agarose electrophoresis of the full length *NCSTN* transcript (single band seen in the patient harbouring the mutation and a control subject) or upon sequencing the full length transcript (cDNA relating to the exon 8/ exon 9 boundary shown in Figure 5.4). There was no significant difference in *NCSTN* mRNA expression in whole blood lymphoblasts taken from the patient harbouring the mutation versus healthy volunteers. Whilst computational splice site analysis predicts that this variant may affect splicing, there is no experimental evidence to suggest that it is pathogenic. This variant is therefore of uncertain pathogenicity.

This lady had severe HS in the axillae, groin and particularly the buttocks comprising painful nodules, cysts, abscesses, sinus tracts and scarring (Figure 5.1b). She was an active smoker with a BMI of 37. Her disease failed to respond to treatment with clindamycin and ciprofloxacin, however some flares partially responded to treatment with Augmentin. Azathiaprine and oral steroids were of some benefit.



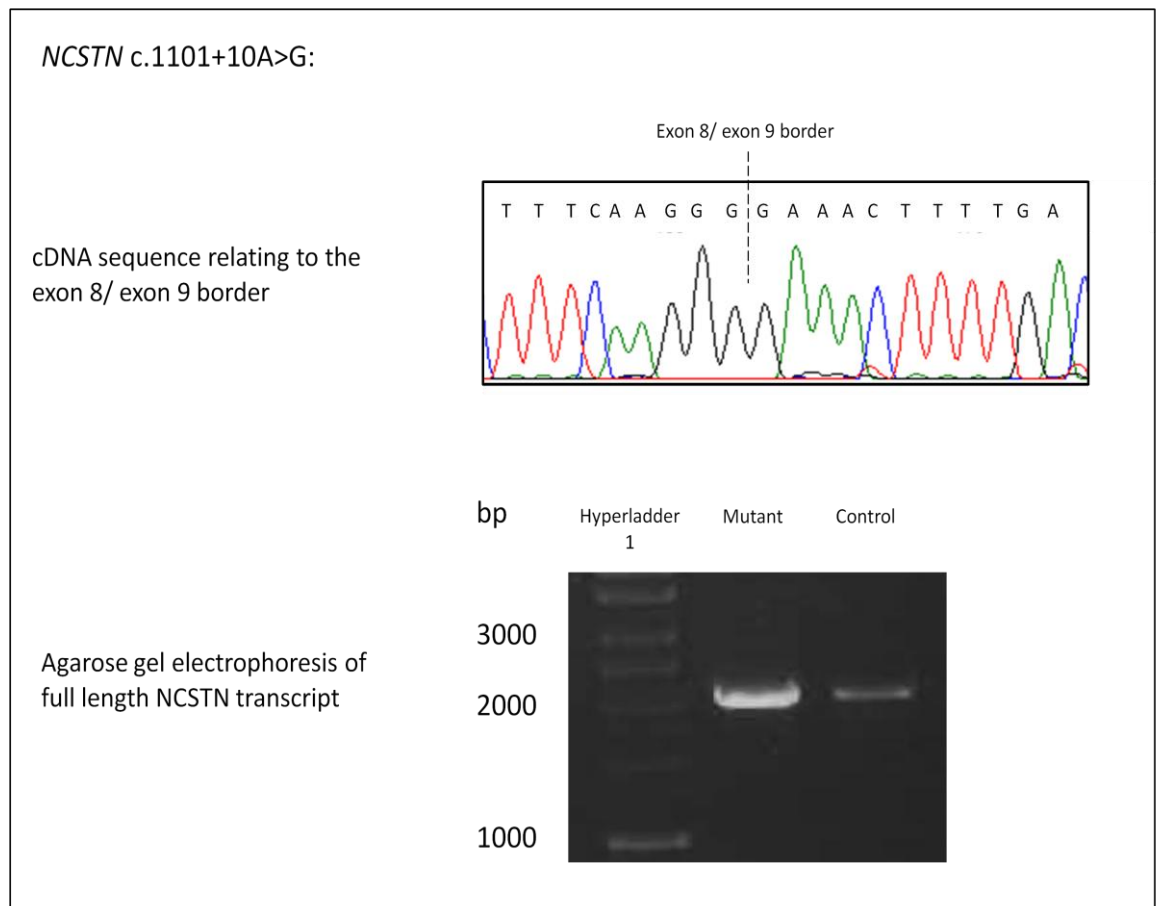
**Figure 5.4. Sequence and agarose electrophoresis of the full length *NCSTN* transcript in the individual harbouring an *NCSTN* c.996+7 G>A variant.** No abnormality was detected in the cDNA transcript, the region relating to the exon 9/ exon 10 border was normal. No extra bands were detected on agarose electrophoresis of the full length *NCSTN* transcript to imply that the variant affects splicing.



## **5.6 Identification of a novel heterozygous substitution in the *NCSTN* exon 9/ intron 9 splice site of *NCSTN* (c.1101+10 A>G)**

A second novel heterozygous splice site substitution was detected in the donor splice site of *NCSTN* exon 9 (c.1101+10A>G) in a 24 year old male with no family history of HS (Figure 5.1c). The variant was absent in 200 European controls but computational splice site analysis suggests that this mutation is unlikely to have a significant effect on splicing (ASSED server<sup>278</sup>). RNA was extracted from whole blood lymphoblasts and converted to cDNA as previously described. Agarose electrophoresis and sequencing of the full length *NCSTN* transcript did not support the presence of a mutant transcript (Figure 5.5). One band was seen on the agarose gel, as was the case with a healthy volunteer control subject, and there were no sequence anomalies detected (cDNA sequence related to the exon 9/ exon 10 boundary shown in Figure 5.5). *NCSTN* transcript abundance within whole blood lymphoblasts was within the range of healthy volunteers. This variant is therefore unlikely to be pathogenic.

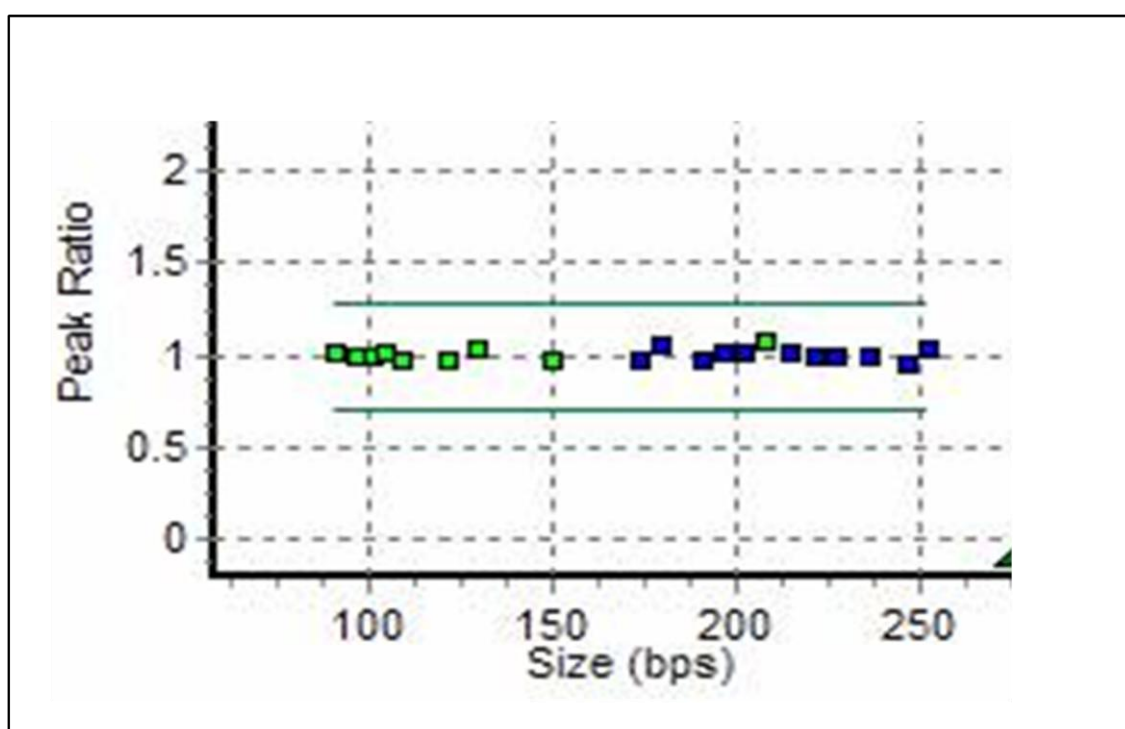
This gentleman had painful nodules, cysts, abscesses, sinus tracts and scarring in the axillae, groin, buttocks and genital region. He was an active smoker with a BMI of 24. His disease failed to respond to numerous antibiotics including oxytetracycline, clindamycin and rifampicin but surgical excision of areas over his buttocks was of some benefit.



**Figure 5.5. Sequence and agarose electrophoresis of the full length *NCSTN* transcript in the individual with an *NCSTN* c.1101+10 A>G variant.** No abnormality was detected in the cDNA transcript and the region relating to the exon 8/ exon 9 border was normal. No extra bands were detected on agarose electrophoresis of the full length *NCSTN* transcript to imply that the variant affects splicing.

## 5.7 No large scale deletions or duplications detected in *NCSTN*, *PSENEN* and *PSEN1*

DNA from all 48 subjects was analysed for whole exon or gene deletions or duplications in *NCSTN*, *PSENEN* and *PSEN1* by MLPA. Probes were designed to target all exons of *NCSTN*, *PSENEN* and *PSEN1*. No copy number variation was detected in this cohort. A representative plot of normalised peak ratios is shown in Figure 5.6.



**Figure 5.6.** A plot of the normalised MLPA peak ratios derived in one individual with HS (probes designed to assess copy number of the exons of *NCSTN*). Green squares represent probes designed to bind to exons of interest (in this case, relating to *NCSTN*). The blue squares represent control probes. The green lines represent the upper and lower thresholds above and below which results would be interpreted as a duplication or deletion respectively. All probes relating to *NCSTN* and all control probes fall within the normal limits, thus no whole exon or whole gene duplications or deletions were detected.

## **5.8 A pilot study to investigate the role of *NCSTN* mutations in nodulocystic acne vulgaris revealed a novel missense variant (*NCSTN* c.1315 G>A, p.Val439Ile)**

As discussed in section 3.4, it was noteworthy that the individual found to harbour an *NCSTN* c.1125 G>A mutation reported a past history of severe acne vulgaris and had cutaneous features of active nodulocystic acne vulgaris over his neck at the time of examination. In total, 7 of the 8 globally reported *NCSTN* mutations at the time of this study appeared to have nodulocystic acne vulgaris/ acne conglobate over their neck, chest or back. We therefore hypothesised that mutations in *NCSTN* underlie nodulocystic acne.

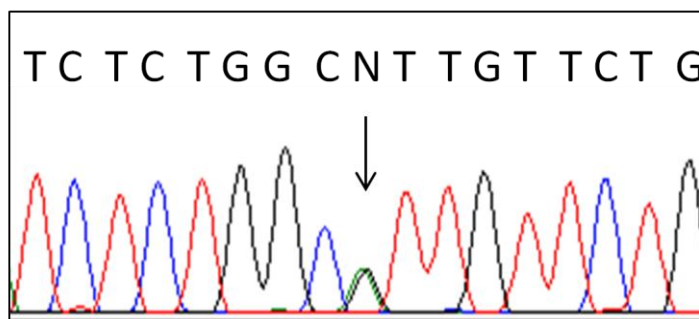
### **5.8.1 Patient cohort**

Forty eight patients with nodulocystic acne vulgaris were identified, consented and recruited from general Dermatology clinics at St. John's Institute of Dermatology. The diagnosis was confirmed by Dermatologists. The coding regions and associated splice sites of all of the exons of *NCSTN* in which mutations had been reported in HS (exons 3, 4, 5, 8, 9, 11, 13, 15) were amplified by PCR using exon-flanking intronic primers and Sanger sequenced.

### **5.8.2 A novel missense variant (*NCSTN* C.1315 G>A, p.Val439Ile) was detected in one individual with nodulocystic acne vulgaris**

A novel heterozygous missense variant in *NCSTN* exon 11 (c.1315 G>A, p.Val439Ile) was identified in a 21 year old female (Figure 5.7). She reported no history of HS and no family history of HS or nodulocystic acne. The variant was not observed in 200 control subjects of European origin. The missense variant is predicted to lead to the substitution of an evolutionary conserved valine residue with an isoleucine residue (Figure 5.8). It is predicted to be tolerated on SIFT (score 0.32) and probably damaging

on PolyPhen-2 (score 0.999). This variant is therefore of uncertain pathogenicity. No other variants with a population frequency of less than 1% were detected in any other individuals within the cohort.



**Figure 5.7. A novel heterozygous substitution in *NCSTN* in a patient with nodulocystic acne vulgaris.** The *NCSTN* c.1315 G>A variant in exon 11 is indicated by an arrow.

Human	- QSQPLPPSSLQRF LRA - RNISGVVLADHSGAFHNK - YYQSIYDTAEN
Mole rat	- QSHALPPSSLQRF LRA - QNISGVVLADHND SFHNL - YYQSIYDTAEN
Guinea pig	- SSHALPPSSLQRF LRA - RNISGVVLADHND SFHNL - YYQSIYDTAEN
Giant panda	- QSQPLPPSSLQRF LRA - RNISGVVLADHSAVFHNH - YYQSIYDTAEN
Pig	- QTQPLPPSSLQRF LRA - RNISGVVLADHSTVFHNH - YYQSVYDTAEN
Horse	- YSKPLPPSSLQRF LRA - RNISGVVLADHLNVFHNH - YYQSIYDTAEN
Dog	- QSQPLPPSSLQRF LRA - RNISGVVLADHSTVFHNH - YYQSIYDTAES
Little brown rat	- QSQPLPPSSLQRF LRA - RNISGVVLTDHATVFHNH - YYHSVYDTAEN
Mouse	- QSQALPPSSLQRF LRA - RNISGVVLADHSGSFHNH - YYQSIYDTAEN
Rabbit	- ESQALPPSSLQRF LRA - RNISGVVLADHSGSFHNK - YYESIYDTAEN
Rat	- QSQALPPSSLQRF LRA - RNISGVVLADHSGSFHNH - YYQSIYDTAEN
Rabbit	- ESQALPPSSLQRF LRA - RNISGVVLADHSGSFHNK - YYESIYDTAEN
Cow	- QSQPLPPSSLQRF LRA - RNVSGVVLADHSD SFHNL - YYQSIYDTAEN
Chinese hamster	- QSQALPPSSLQRF LRA - RNISGVVLADHSGSFHNH - YYQSIYDTAEN
Opossum	- PPRS LPPSSLQRF LRG - QNISGVVLTDYSHSFSNR - YYQSIYDTPEN
Tasmanian devil	- PQNS LPPSSLQRF LRS - QNISGVVLTDHPGSFYNR - YYQSIYDTPEN
Turkey	- FSQSLPPSSFQRF LRA - RHIPGVVLTDHQASFQNR - YYQSMYDTPEN
Chicken	- FSQPLPPSSFQRF LRA - RHIPGVVLSDHQASFQNR - YYQSMYDTPEN
Finch	- QSQPLPPSALQRF LRAR RGLPGVLLADHRGAFNR - YFQSIYDTAES
Frog	- RSQPLPPASLQRF LRV - RNIPGVVLTDHRTAYSNR - YYHSVYDTADN
Elephant	- QSQPLPPSSLQRF LRA - RNISGVVLADHSAAFHNH - LLTS - - - - -
Elephant	- VSQPLPPSSFQRF LRV - RPIGLVLT DHEKA FNH - YYESMYDDAEN
Zebrafish	- VSQPLPPSSFQRF LRV - RPIGLVLT DHEKA FNH - YYESMYDDAEN
Tilapia nilotica	- FSQPLPPSSFQRF LRA - RPIGVV IQNHN SSFTNK - YYESMYDNAEY
Stickleback	- FSQPLPPSSFQRF LRA - RPIGVV IQNHN SSFTNK - YYESMYDNAEY
Pufferfish	- FSQPLPPSSFQRF LRA - RPIGVV IQDHQSSFTNRWFY ES MYDNAGY

**Figure 5.8. Conservation of the *NCSTN* Val439 residue.** The amino acid is conserved through mammals but is replaced by a Leucine residue in Zebrafish.

## 5.9 Discussion

At the time of study, gamma-secretase mutations were known to underlie some familial forms of HS. The aim of this study was to investigate how prevalent gamma-secretase mutations were in the general disease population. This was important both in terms of clarifying the genetic architecture of the condition and providing perspective on the translational potential of this discovery. The study cohort was recruited from a tertiary referral HS clinic and their demographic details closely matched those of previously reported HS cohorts. There was a female predominance, average age of onset was 21.4, raised average BMI, increased likelihood of smoking and 42% reported a family history of the condition. The average Sartorius and DLQI scores were high at 49.9 and 14.87. These measures of disease severity highlight one potential limitation of this study which was that individuals recruited from a tertiary referral clinic are unlikely to be truly representative of the general disease population.

Three of forty eight individuals harboured novel heterozygous variants in *NCSTN*. No other variants with a population frequency <1% or CNV was detected in *NCSTN*, *PSEN1* or *PSENEN*. The three novel variants comprised a missense change and two splice site substitutions. Whilst it is relatively easy to determine the prevalence of rare variants in a gene, this study has demonstrated that it is somewhat harder to determine their pathogenic potential. The *NCSTN* c.996+7 G>A splice site substitution was predicted to affect splicing using splice site prediction software however there was no evidence to support the presence of a mutant transcript and overall transcript abundance appeared unaltered. The second splice site substitution (c.1101+10 A>G) was ten bases downstream of the *NCSTN* exon9/ intron 9 donor splice site and did not appear to affect splicing or overall transcript abundance. The missense substitution (c.553 G>A) affected a poorly conserved amino acid, was not predicted to be damaging on SIFT or Polyphen-2 and did not confer any effect on splicing. A maximum of three individuals

therefore appear to harbour potentially pathogenic variants in *NCSTN*, *PSENEN* or *PSEN1* in this cohort, representing <7% of the group. This figure is likely to be inflated given that there is insufficient evidence to confirm the pathogenicity of any of these variants. Two further studies have since screened smaller cohorts of patients for mutations in all six of the gamma-secretase genes (*NCSTN*, *PSENEN*, *PSEN1*, *PSEN2*, *APH1A*, *APH1B*). No rare variants (population frequency < 1%) were detected in a Welsh cohort of 20 individuals<sup>279</sup>. Only one of nine Japanese patients was found to harbour rare variation in those genes, a heterozygous missense substitution in *PSEN2* (p.Thr421Met), but that variant was deemed non-pathogenic<sup>141</sup>. Taken together, these studies suggest that gamma-secretase mutations in *NCSTN*, *PSENEN* and *PSEN1* underlie a minority of HS cases, ranging from 0 - 6.3%. There is no current evidence to suggest that variation in *PSEN2*, *APH1A* or *APH1B* underlies HS. Whilst our study represents the largest patient cohort to be studied to date, the numbers are still relatively small and to gain a more accurate measure of mutation prevalence one would need to screen a much larger number of individuals.

Around 40% of HS patients report a family history of the condition, the majority consistent with autosomal dominant inheritance, yet less than 7% of cases appear to occur as a consequence of mutations in *NCSTN*, *PSENEN* and *PSEN1*<sup>22</sup>. The genetic basis of the remaining familial cases (representing around 33% of the disease population) and apparently sporadic cases therefore remains unclear. Whilst it is possible that coding variation may have been missed or that there was pathogenic variation in unscreened non-coding regions (promoter regions, enhancer regions and intronic variation for example) it would seem unlikely that these explanations account for the vast majority of disease cases. An alternative theory is that those genes do not underlie the development of HS in the remaining cases and that there is further genetic heterogeneity underlying HS.

One of the individuals reported in Chapter 3 (section 3.4) who harboured a mutation in *NCSTN* (c.1125+1 G>A) presented with features consistent with nodulocystic acne vulgaris over their neck. At least 7 of 8 globally reported *NCSTN* mutations were also associated with a nodulo-cystic acne vulgaris/ acne-conglobate-like presentation over the neck, chest and back in addition to the more characteristic features of HS<sup>135-138</sup>. A pilot study was therefore performed to examine whether any of 48 individuals with nodulocystic acne harboured mutations in any of the exons of *NCSTN* in which mutations had been reported in HS. Only one variant with a population frequency < 1% was detected in this cohort (c.1315 G>A, p.Val439Ile). This variant was of uncertain pathogenicity. A role for *NCSTN* in nodulo-cystic acne cannot therefore be excluded on the basis of this work, particularly given that only select exons of *NCSTN* were sequenced. The nature of the identified variant was not sufficient to justify pursuing this work further. A larger cohort of individuals would need to be screened for variants in all exons of *NCSTN* to formally validate or exclude involvement of that gene in the pathogenesis of nodulocystic acne vulgaris.

In summary, the results of this study suggest that gamma-secretase gene mutations underlie a minority of HS cases and that there is likely to be further genetic heterogeneity underlying HS. This project therefore re-focussed on genetically investigating familial cases of HS with no identifiable gamma-secretase mutations with a view to discovering further disease-causing genes.



## **Chapter 6: Investigation to determine further disease causing genes in HS**

### **6.1 Introduction**

#### **6.1.1 Hypothesis and objectives**

Gamma-secretase gene mutations appear to underlie a minority of HS cases, implying that there is likely further genetic heterogeneity in HS. It was therefore hypothesised that further and as yet unidentified genes underlie some cases of HS.

The aims of this study were two-fold:

- 1) Investigate large multiplex kindreds, preferentially those in which gamma-secretase gene mutations had been excluded, for novel disease-causing variants.
- 2) Investigate meticulously phenotyped subgroups of unrelated individuals with a family history of HS (demonstrating autosomal dominant inheritance, preferentially those in which gamma-secretase gene mutations had been excluded) for novel disease-causing variants.

#### **6.1.2 Methodological considerations**

The methods employed to identify disease-causing variants in Mendelian disorders have rapidly evolved over the last 40 years, gathering significant momentum in the last 5 years with the advent of next generation sequencing. The introduction of DNA recombinant technology and polymerase chain reaction (PCR) in the 1980's enabled the identification of an increasing number of disease genes via positional cloning. This involved identifying the approximate chromosomal location of a gene via linkage analysis, mapping out the genes in the region by DNA cloning and screening the genes in affected individuals for mutations. This was a long and technically challenging

process. The introduction of genome-wide genotyping technology has facilitated faster and more accurate linkage analysis. Large scale genome projects have provided increasingly more accurate and comprehensive marker maps, gene identity and location data and specific sequence data, negating the need for labour intensive DNA cloning. Furthermore, the efficiency of direct sequencing for mutation detection significantly improved following the development of Sanger sequencing. All of these technological advancements have made positional cloning easier and more efficient, resulting in an explosion of papers documenting new disease variants. In 2009, the utilisation of next generation sequencing led to a dynamic shift in sequencing capability and has been used alone, and in combination with linkage data, to great effect in determining disease-causing variants in Mendelian disease.

#### **6.1.2.1 Linkage analysis**

Genetic mapping using linkage analysis is a well established technique for identifying the approximate chromosomal location of disease causing variants. The principle of this technique is that when maternal and paternal chromosomes pair up in germ cells there is a cross-over of genomic material. This process is termed recombination. The further apart two genes are on a chromosome, the more likely it is that they will cross-over and end up on different chromatids. Conversely, the closer two genes are together, the less likely it is that cross-over will occur. As a disease-causing variant is passed down through a family, the original flanking haplotype on which the mutation resides will steadily become shorter and shorter as recombinations occur. Methodologies to ascertain when and where recombinations have occurred in a pedigree are based on genotyping respective family members. Numerous techniques have been employed in the past to generate this data, including the use of microsatellite markers, but the recent development of SNP genotyping microarrays has facilitated rapid analysis of millions of SNPs scattered throughout the genome. These SNPs act as genetic markers. Every

disease gene will be flanked by two of these markers unless the mutation is located in a distal region of a chromosome (the exact distance of the relevant markers from the disease gene will vary). Thus by genotyping sites of variation in all family members it is possible to determine regions of maternal and paternal DNA in individuals and calculate where re-combinations have occurred. Recombinations between the disease gene and its flanking markers are highly unlikely because of the sheer proximity of these markers. Regions such as this can be mapped, thus highlighting candidate regions within the genome. The principle of linkage analysis is that the disease allele, situated within only one of those regions, will co-segregate with disease status in any given pedigree. The size of the candidate region will vary depending on the size and structure of the pedigree studied.

Genotype data is analysed using linkage analysis software. Information required for accurate analysis is the exact pedigree structure, disease status of the individuals involved, mode of inheritance (eg. autosomal dominant), estimated population frequency of a disease causing variant, the estimated genetic distance between variants on the SNP genotyping microarray and the genotype data for the individuals involved. Variable analyses can be performed depending upon the nature of the disease in question. For example, HS can occur at any age thus it is impossible to conclude that an individual is definitely unaffected. This non-penetrance could be modelled or an affected-only analysis may be appropriate in that instance. Linkage analysis generates a set of data for each pedigree demonstrating the likelihood of linkage at any given point in the genome. This data comprises a sequence of logarithm of odds (LOD) scores. LOD scores represent the likelihood of generating the observed experimental data if markers are indeed linked versus the likelihood of achieving such data if they were not linked. It is a logarithmic scale whereby positive scores support the presence of linkage and negative scores do not. A LOD score of 3 means that there is only a 1:20 chance

that there is no linkage at that site and is therefore taken as good evidence for linkage. Another way to interpret such linkage data is that linkage cannot be excluded in any regions with a LOD score above -2.

#### **6.1.2.2 Whole exome sequencing**

Whole exome sequencing (WES) is a next generation sequencing method that involves sequencing all of the annotated protein-coding regions of the human genome. It is a fast, efficient and, as prices have fallen, increasingly economical way to study the protein coding regions of the human genome. Since it was first introduced and successfully implemented in 2009<sup>280</sup> it has revolutionised the genetic investigation of Mendelian disorders and exponential numbers of papers utilising this technology have emerged over the last three years (>400). The principle method involved is that DNA is fragmented, amplified and hybridised to oligonucleotide baits that select out regions of the exome (exome capture). The bound DNA is selected out (representing exomic regions), eluted from the baits, washed, further amplified and then sequenced using massively parallel next generation sequencing. This generates short sequences which are mapped to the reference human genome and aligned. Non-exomic regions are filtered out and then variants can be identified and called. Variants are subsequently annotated and compiled in a database for bioinformatic analysis.

##### ***6.1.2.2.1 Genetic diseases investigated using WES***

Genetic interrogation of recessive traits has always been relatively successful due to the limited number of novel or very rare homozygous variants present in any genome. Consequently, when combined with accurate linkage data (or even in the absence of such data) and the numerous databases providing the population frequency of any genetic variation, WES has proven itself to be a successful tool in determining disease-causing variants in recessive diseases<sup>281; 282</sup>.

Analysing dominant conditions, such as HS, has always been more challenging. Due to the comparatively large numbers of novel and rare heterozygous variants present in any individual, the isolation of a disease-causing variant from a WES dataset requires as much additional support as is available. This may comprise good linkage data from large pedigrees with multiple affected members, access to a reliable control cohort against which the frequency of variation in any one gene can be assessed and access to databases stating the population frequency of rare variants. Familial analysis (assuming large and informative pedigrees are available) confers a benefit over analysing unrelated subjects. Multiple and far removed individuals from any one family can be subjected to exome sequencing. The validity of any shared candidate variants can be confirmed by performing linkage analysis and assessing for co-segregation of any identified variants with disease status. When studying small numbers of unrelated cases in an apparently genetically heterogeneous condition the diagnostic accuracy, and thus the exact phenotype being studied, is paramount. There have been recent successes in studies that have both used pedigrees and unrelated cases<sup>283; 284</sup>.

Even non-inheritable diseases (those that affect reproductive capability to the extent that they cannot be passed on) have been effectively interrogated using WES. By studying apparently sporadic cases and their parents it is possible to determine if variants are truly de novo. When there are enough individuals, the number of de novo variants can be analysed and mapped to a common disease gene. This principle was effectively employed in investigating the genetic predisposition to autism<sup>285</sup>

The role of WES in determining the genetic architecture of complex diseases is yet to be determined. Whilst WES can identify rare coding variants, many thousands of individuals would be required to implicate any such variants in disease pathogenesis.

### **6.1.2.3 Potential disadvantages of WES**

Whilst whole exome sequencing provides a very powerful and increasingly utilised platform upon which to interrogate genetic disease, there are some disadvantages and limitations to this form of sequencing. It can only target known genes, thus any protein-coding regions of the genome that are not yet known will not be covered using this technique. The aim of WES is to target the “functional” regions of the genome, however it is becoming increasingly evident that non-coding regions play a significant role in transcription, splicing, translation and thus protein expression. These comprise gene promoters, enhancers, silencers, regions that affect splicing (such as branch sites) and non-coding RNAs. Whilst an alternative method, such as whole genome sequencing, may circumnavigate this problem, it would not alter the fact that the functional impact of many of these non-coding regions is unknown and thus the interpretation of any such data is extremely challenging. Furthermore, whole genome sequencing is up to five times more expensive and generates at least five-fold more data, posing significant problems regarding data storage and analysis. A further limitation of WES is that it is not a proven way of detecting structural variation such as copy number variation (CNV) or translocations.

As with any experiment, there will always be intra- and inter-run variability. For example the read-depth and sequencing quality will vary across the target exome and will not be consistent between individuals. This, compounded by any potential variability in the bioinformatic pipeline used to generate the results, may result in false-positive and false-negative variant calls and significantly affect experimental outcomes. In the context of HS, if two individuals from one pedigree are exome sequenced it may be that some shared variants are missed because of false-negative calls. A further consideration is the heritability of the condition in question. Even with a low read-depth, genuine homozygous variants are still likely to be identified. With heterozygous

conditions, such as HS, the risk of false-positive and false-negative calls is much greater. A high coverage depth is therefore an essential requirement when studying dominant conditions. Another potential variable in data analysis is the accuracy of read-alignment. Each of the sequences generated via massively parallel sequencing in WES have to be aligned to their corresponding sequence in the genome. The repetitive nature of large parts of the human genome make this difficult and thus the filtering strategies employed to filter out misaligned reads are paramount to the success of the technique.

## **6.2 Genetic methods and patient selection**

### **6.2.1 WES, combined with linkage analysis in familial cases, was deemed an appropriate method by which to further interrogate the genetic architecture of HS**

Despite the above mentioned disadvantages, WES remains an extremely powerful, fast and relatively cost-efficient way to analyse the functional parts of the human genome and experimental techniques and bioinformatic analysis methods have significantly improved since their introduction in 2009. Our group set up WES at King's College London and has successfully used the technique to identify multiple novel disease causing mutations over the last 3 years<sup>281; 283; 284; 286</sup>. The advantages of this cutting edge technology, combined with the experience in the group, meant that it was a highly appropriate experimental approach to employ in searching for new disease-causing mutations and genes in HS. Linkage analysis was performed as an adjunct to WES in multiplex kindreds to facilitate more effective WES data filtering.

### **6.2.2 Whole exome sequencing**

#### **6.2.2.1 Selection of individuals for exome sequencing**

As layed out in the Chapter aims, there were two broad groups of patients analysed as part of this study:

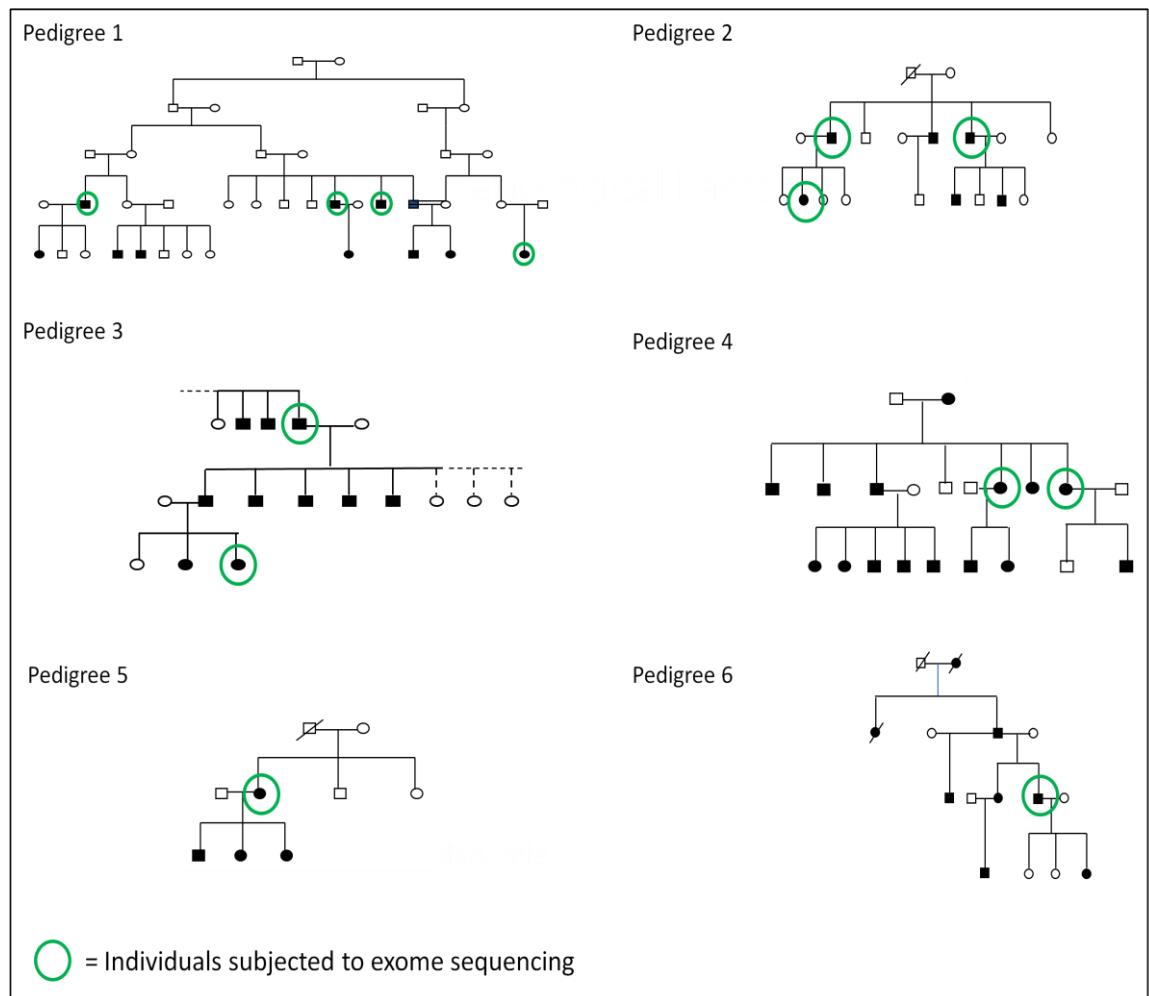
1/ Related individuals: Pedigree-centred analysis of individuals from six large multiplex kindreds

2/ Unrelated individuals: Phenotype-centred analysis of phenotypically grouped unrelated individuals



#### **6.2.2.1.1 Related individuals**

Six pedigrees were identified that demonstrated autosomal dominant inheritance and from whom at least one individual had been screened to exclude novel or rare (population frequency <1% in both the 1000 Genomes Project database and the NHLBI Exome Sequencing Project EVS) variation in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1*. Five of these pedigrees had been studied in Chapter 3, all six are shown in Figure 6.1. The number of individuals subjected to WES in each pedigree was decided by analysing pedigree structures to determine which individuals would provide the most contrasting and therefore revealing genetic information (dependent upon availability of relevant clinical samples). For example, the two individuals sequenced in pedigree 3 represent the two most genetically distant (2 meioses) subjects that were available within that pedigree. One would therefore expect fewer shared variants upon comparing results generated from both individuals than if father and daughter (1 meiosis) were compared for example, thus aiding analysis. A total of 13 individuals from the six pedigrees were subjected to exome sequencing.



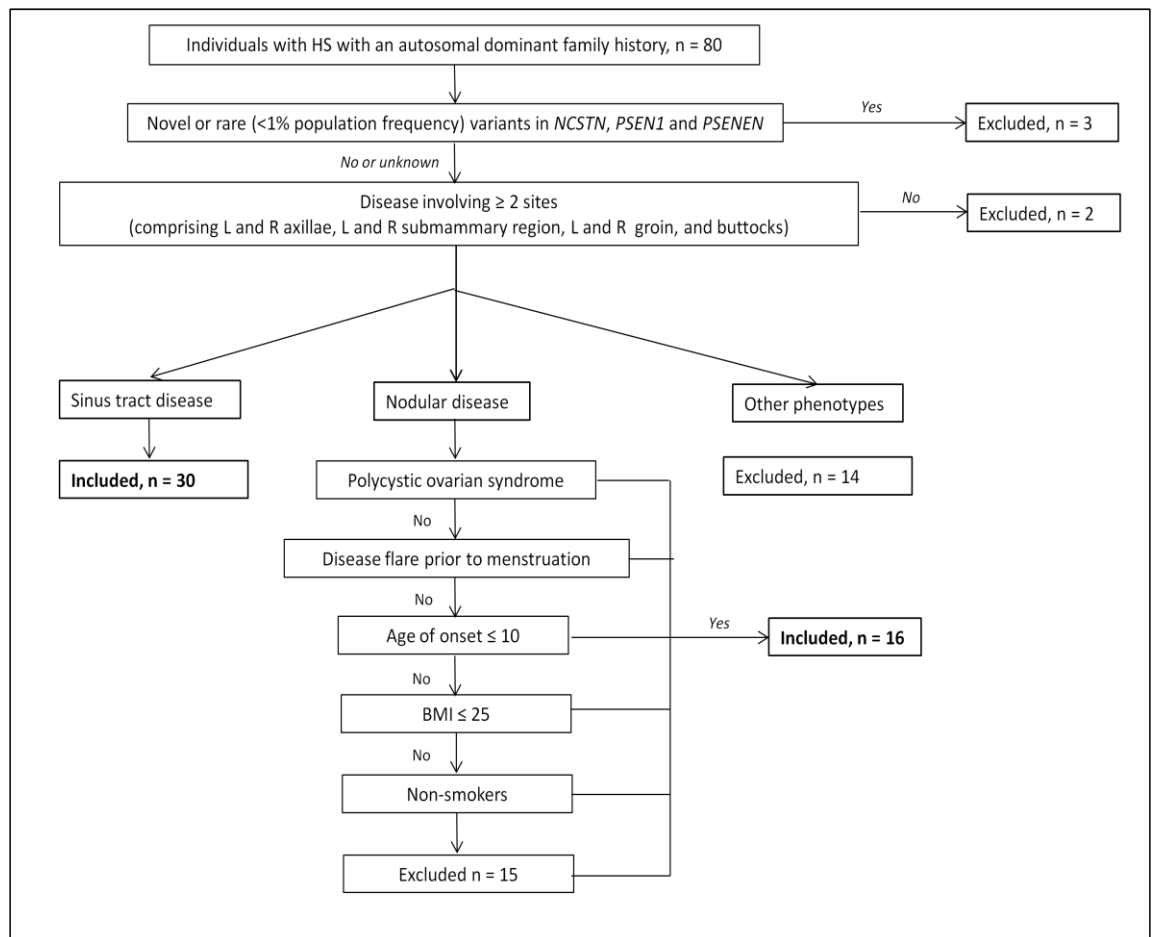
**Figure 6.1. Multiplex kindreds studied using WES.** All individuals circled were subjected to WES.

#### 6.2.2.1.2 *Unrelated individuals*

A further 53 unrelated probands, all but three with a family history of HS, were chosen based upon their clinical phenotype. When this study was designed there was no phenotypic classification scheme for HS. By taking detailed clinical histories and performing meticulous clinical examinations patients were categorised into 4 broad groups. The first comprised a small group of individuals with severe cutaneous disease (nodules, multiple sinus tracts and significant genital oedema) and an associated inflammatory arthritis (exacerbations corresponded with cutaneous flares, variable sites including spondyloarthropathy, sacro-ileitis and large joint oligoarthritis). Four of these

seven individuals reported a family history of HS and three were apparently sporadic cases. Due to the rare but entirely consistent signs and symptoms observed in these individuals all seven were taken forward for exome sequencing. The second group had predominantly sinus tract disease in all affected areas (comedones and nodules also often present). The third group comprised females with inflammatory nodules, no sinus tracts and a history suggesting hormonal involvement (disease associated with premenstrual flares, many individuals formally diagnosed with poly-cystic ovarian syndrome). The final group comprised individuals with comedonal and cystic disease and no sinus tracts. These patients were predominantly male and almost exclusively smokers.

The sinus tract disease and nodular disease groups were the most phenotypically distinct and were thus considered for exome sequencing. Individuals with sinus tract disease were chosen ahead of those with nodular disease based on the severity and consistency of that presentation (30 individuals). A further 16 candidates were selected from the nodular disease group based on the algorithm shown in Figure 6.2. This was designed to identify people who were more likely to have a genetic basis for their disease.



**Figure 6.2. The algorithm used to select individuals within the sinus tract and nodular disease cohorts for WES.**

17 of these individuals had previously been screened and not found to harbour any mutations in *NCSTN*, *PSEN1* or *PSENEN* as part of the study described in Chapter 5. The remaining 36 individuals had not however been sequenced, thus raising the question as to whether they should be screened prior to inclusion in the study. Upon considering the time required to complete such sequencing, the projected sequencing costs in relation to the falling cost of WES and the low prevalence of gamma-secretase mutations in the general disease population, it was decided that pre-screening was not essential.

### **6.2.2.1.3 A summary of individuals taken forward for exome sequencing**

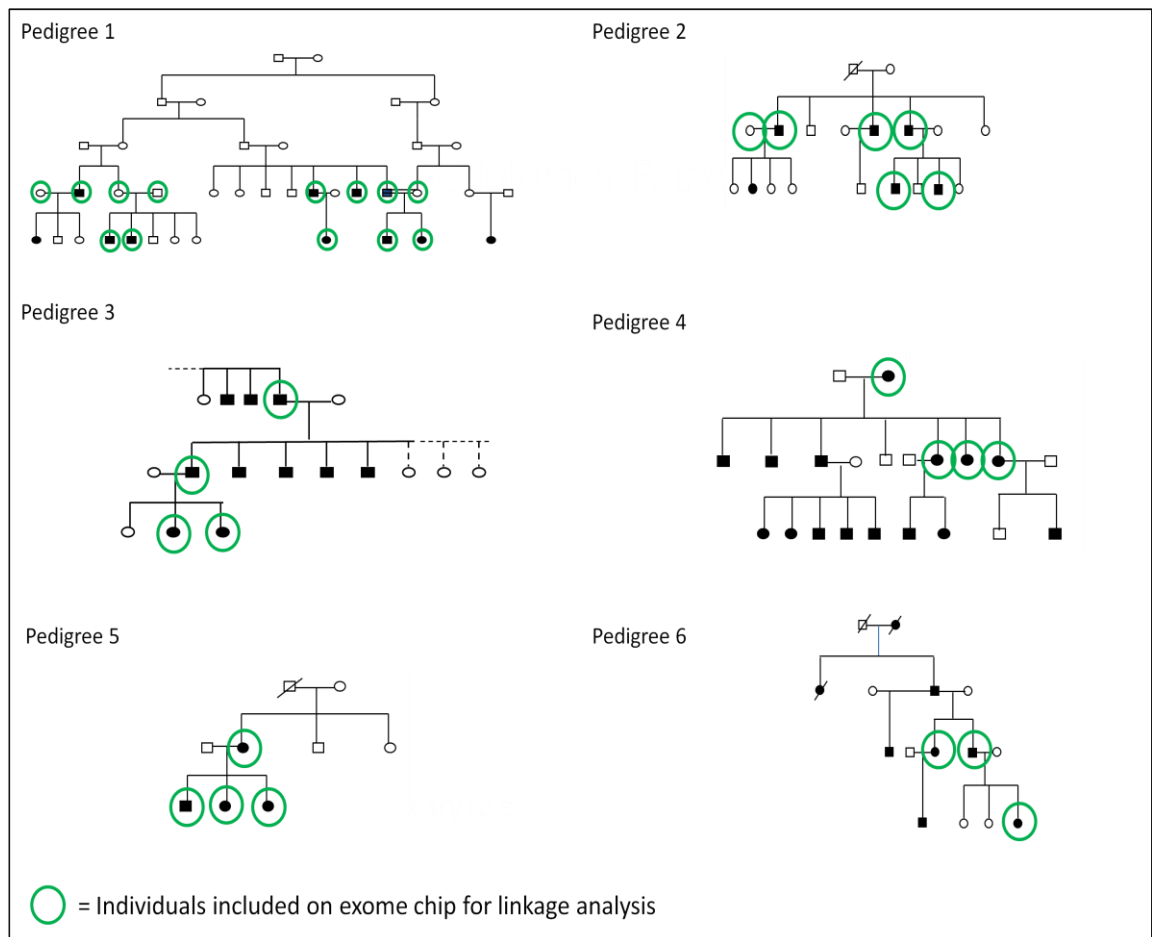
In total, 66 individuals were subjected to WES. These comprised 13 individuals from multiplex kindreds (to be used in pedigree-centred analyses) and 53 unrelated and phenotypically grouped cases (to be used for phenotype-centred analyses). Individuals from two of the multiplex kindreds had sinus tract disease and thus one individual from each pedigree was included in the sinus tract disease cohort as part of the phenotype centred analysis. In total there were 7 individuals with HS/ arthritis, 32 with sinus tract disease and 16 with nodular disease (Table 6.1).

**Table 6.1. A table showing the phenotypic subgroups analysed via WES and the number of individuals within each group.**

Clinical subgroup	No. Individuals for analysis
Familial analysis	13
HS and arthritis	7
Sinus tract disease	32
Nodular hormonal disease	16

### **6.2.3 Genotyping of individuals from the multiplex kindreds and linkage analysis**

In addition to exome sequencing individuals from each of the six pedigrees studied as part of the pedigree-centred analysis (detailed in section 6.2.2.1.1, Figure 6.1), all affected and any important unaffected individuals from those kindreds were also genotyped using the Illumina Infinium HumanExome 12v1.1 chip to generate linkage data (genotyped individuals shown in Figure 6.3). An affected only analysis was performed that assumed an autosomal dominant inheritance model.



**Figure 6.3. A figure showing the individuals within each pedigree that were genotyped using the Illumina Infinium HumanExome 12v1.1 chip to generate linkage data.**

## 6.2.4 Data analysis strategy

All coding variants and variants within 10bp of splice sites were included in the analysis along with any copy number variation affecting coding regions.

### 6.2.4.1 Exclusion of individuals harbouring likely pathogenic variants in the gamma-secretase genes

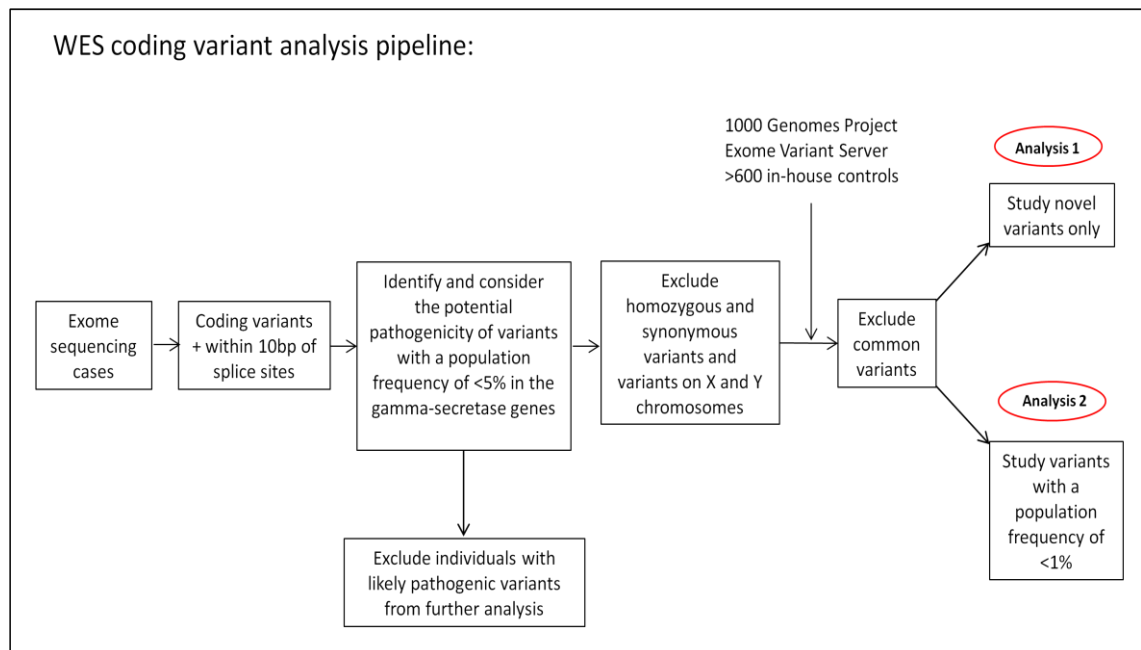
All exomes were initially screened for variants with a population frequency of <5% and CNV in or affecting the gamma-secretase genes *NCSTN*, *PSEN1*, *PSEN2*, *PSENEN*, *APH1A* and *APH1B*. Variants with a population frequency of <5% were defined as having a population frequency <5% in the 1000 Genomes Project database<sup>260</sup>, the

NHLBI Exome Sequencing Project Exome Variant Server (EVS, <http://evs.gs.washington.edu/EVS/>) and in 600 in-house control subjects. All population frequencies stated throughout the remainder of this chapter are based on these criteria.

The potential pathogenicity of each variant was predicted by considering the population frequency, conservation, nature of the variant, predicted functional consequences and where relevant running the variants through pathogenicity prediction software. Missense variants were run through SIFT<sup>276</sup> and/ or Polyphen-2<sup>277</sup> and splice site variants through the Automated Splice Site and Exon Definition Analysis (ASSEDA) server<sup>278</sup>. Any variants that were predicted to be pathogenic were confirmed with Sanger sequencing and the relevant individuals were excluded from further exome sequencing analysis.

#### **6.2.4.2 Analysis of substitutions and small insertions and deletions in individual exomes in which pathogenic gamma-secretase gene mutations had been excluded**

The remaining exome variant profiles were analysed under an autosomal dominant inheritance model. Substitutions and small insertion/ deletions were filtered according to Figure 6.4. Only heterozygous and protein altering (nonsynonymous, splice site substitutions or small coding deletions/ insertions) variants were considered to be pathogenic. As there is no evidence to suggest X or Y linked inheritance in HS, variants on those chromosomes were excluded. Common variation was then excluded by performing two separate analyses (analyses 1 and 2). Only novel variants (variants not reported in 1000 genomes, EVS or >600 in-house controls) were included in the first analysis (analysis 1) and only variants with a population frequency of < 1% were included in the second analysis (analysis 2).

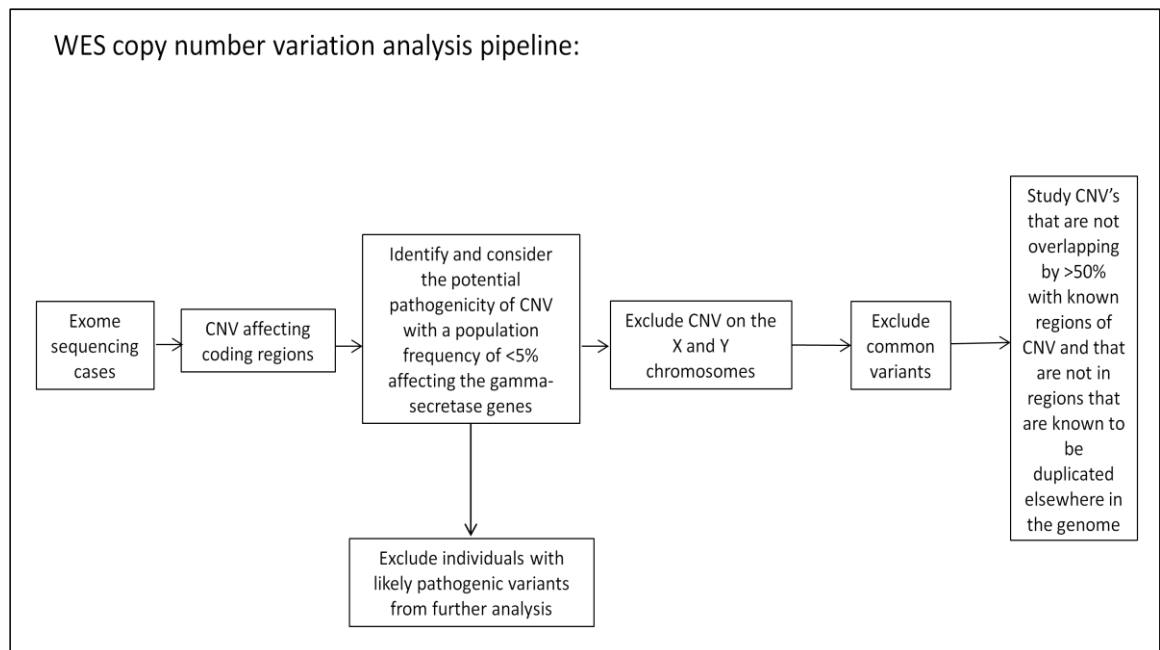


**Figure 6.4. Algorithm for filtering coding variants or variants detected within 10bp of splice site boundaries identified through WES.**

#### **6.2.4.3 Analysis of CNV in individual exomes in which pathogenic gamma-secretase gene variants had been excluded**

Copy number variants were filtered as shown in Figure 6.5. CNV affecting the X or Y chromosomes was excluded. Only CNV's that did not overlap by more than 50% with known regions of common CNV and that did not affect regions known to be duplicated elsewhere in the genome were included for further analysis.





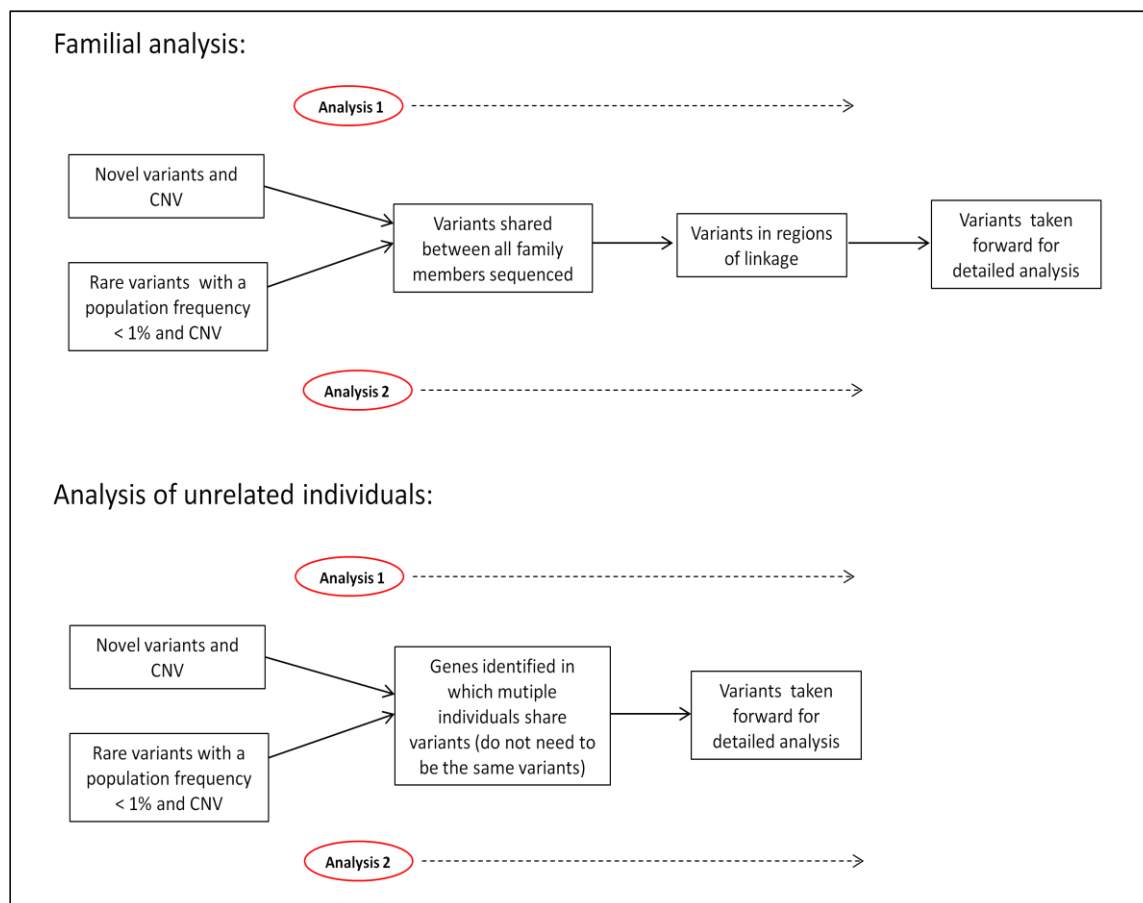
**Figure 6.5. Algorithm for filtering CNV identified by WES.**

#### **6.2.4.4 Combinational analysis of individuals within their respective experimental subgroups**

Filtered coding variants (+ those within 10bp splice sites) and CNV's were then combined for each individual. Novel filtered variants described in Figure 6.4 were combined with the filtered CNV's described in Figure 6.5 for analysis 1. Rare filtered variants (< 1% population frequency) documented in Figure 6.4 were combined with the filtered CNV's described in Figure 6.5 for analysis 2.

Data from the separate individuals was then collated. Regarding the familial group, data from individuals within the same pedigree was pooled. Regarding unrelated individuals, data from individuals within each specific phenotypic group was pooled (Figure 6.6). For familial analysis, any variants shared between all exome sequenced relatives were identified. Only those that fell within defined regions of linkage in the family were taken forward for further study. For the analysis of unrelated individuals, all genes were

identified in which multiple individuals within any one clinical subgroup shared variation. Individuals did not need to share the same variants in these genes.



**Figure 6.6. Algorithm detailing how data from individuals within the familial and unrelated cohorts was generated, collated and filtered prior to ongoing analysis.**

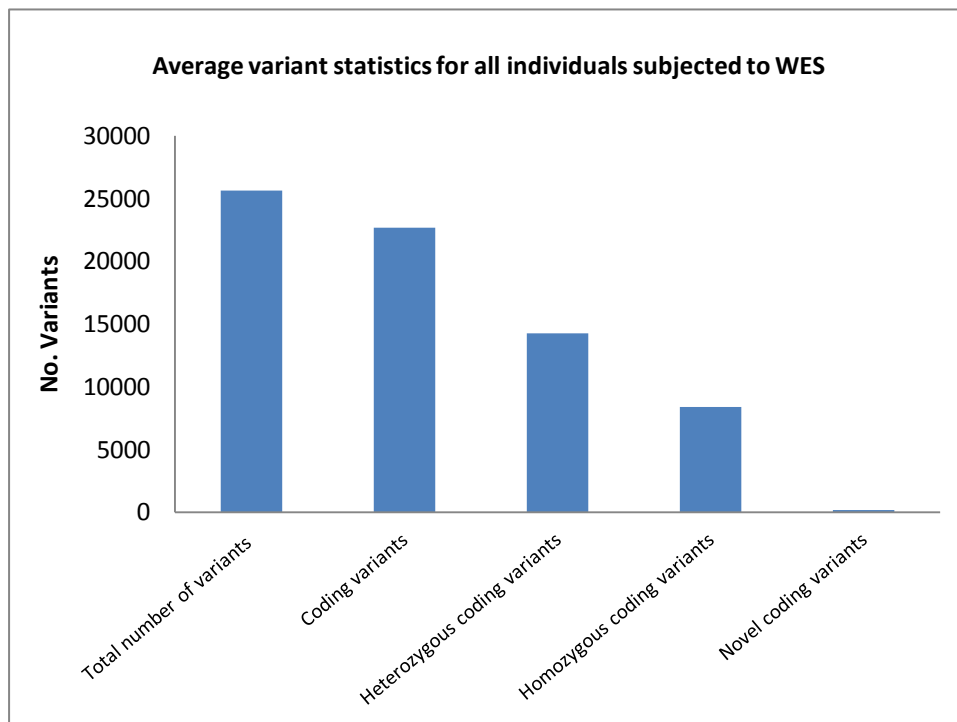
## 6.3 Results

### 6.3.1 Exome sequencing coverage

Across all exomes, 82.84% of reads mapped to target regions (+/- 150 base pairs), 89.25% of the exomes had a read-depth >20. An average of 22657 coding variants were called per individual of which 170 were novel (defined as not present on the EVS, 1000 genomes server or in > 600 in-house control exomes). Detailed coverage statistics and variant profiles are shown in Table 6.2 and Figure 6.7 respectively.

**Table 6.2. Coverage statistics for all individuals subjected to WES**

	<b>Average</b>	<b>Range</b>
Total reads	84655696	40954786 – 136309125
Mapped to target/ %	73.87	69.1 – 83.55
Mapped to target + 150bp/ %	82.84	77.83 – 92.86
Mean coverage	111.97	56.73 – 183.64
≥ 20 x read-depth/ %	89.25	67.9 – 93.98



**Figure 6.7. Variant statistics for all individuals subjected to WES.** The figures represent the averages across all sequenced individuals. Novel variants represent a minority of the total number of variants detected in any individual.

### 6.3.2 Analysis of the gamma-secretase genes in all individuals subjected to exome sequencing

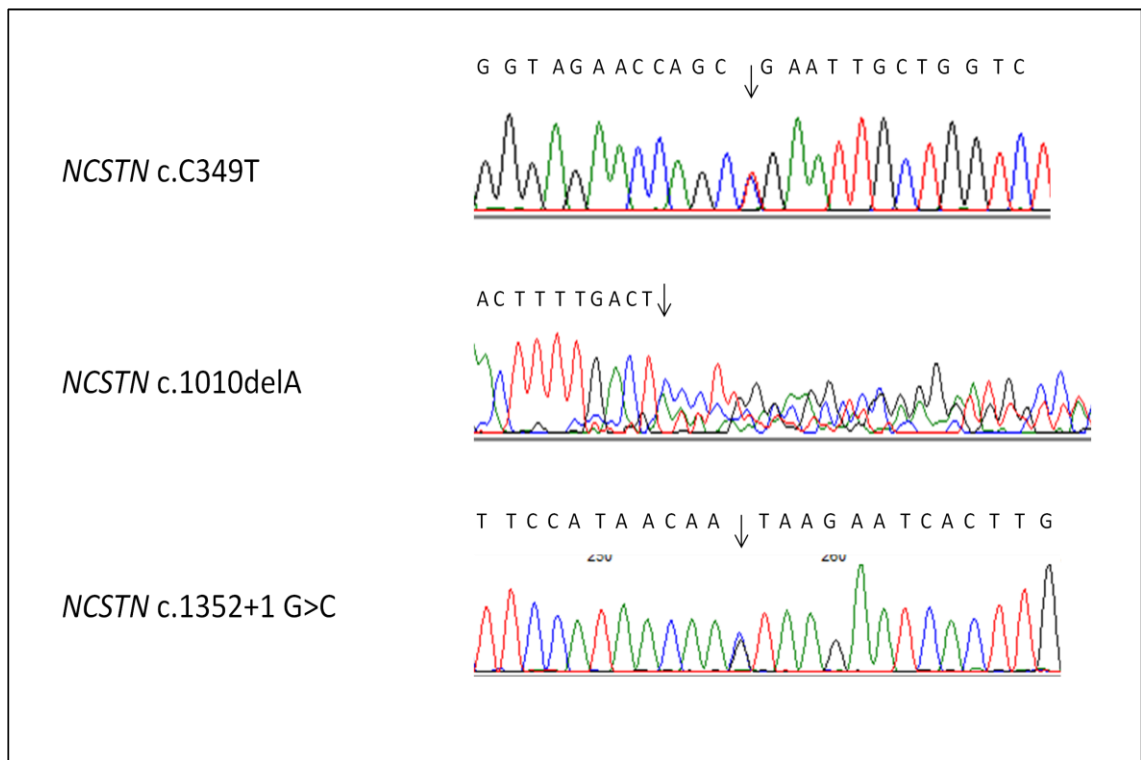
Mutations in *NCSTN*, *PSEN1* and *PSENEN* are known to underlie some cases of HS and on a candidate basis, mutations in the other gamma-secretase genes, *PSEN2*, *APH1A*, *APH1B* may also underlie HS. As a consequence, all exome-sequenced individuals were assessed for variants (synonymous, non-synonymous, splice site, small insertion/deletion variants and copy number variation) in the gamma-secretase genes *NCSTN*, *PSEN1*, *PSEN2*, *PSENEN*, *APH1A* and *APH1B* with a population frequency of less than 5%. All such variants are shown in Table 6.3. No CNV was detected in these genes.

**Table 6.3. All variants with a population frequency < 5% detected in the gamma-secretase genes in the cohort of individuals subjected to WES.**

<b>Gene</b>	<b>Variant</b>	<b>Variant type</b>	<b>Population frequency/ %</b>
APH1A	c.C735G, p.C245W	Nonsynonymous SNV	0.33 - 0.62
APH1B	c.285-7 A>T	Splicing	0 - 2.5
APH1B	c.G642A, p.A214A	Synonymous SNV	0.038 – 0.15
APH1B	c.G472T, p.V158L	Nonsynonymous SNV	0.97 - 2.67
NCSTN	c.T201C, p.S67S	Synonymous SNV	0.33 – 1.61
NCSTN	c.G231C, p.E77D	Nonsynonymous SNV	0.16 - 1.57
NCSTN	c.G237A, p.E79E	Synonymous SNV	0 – 2.33
NCSTN	c.A1083G, p.S361S	Synonymous SNV	0.66 – 1.86
NCSTN	c.C1314T, P.G438G	Synonymous SNV	0.33 - 0.89
NCSTN	c.1352+1 G>C	Splicing	Novel
NCSTN	c.1010delA, p.Y337fs	Frameshift deletion	Novel
NCSTN	c.C349T, p.R117X	Nonsense substitution	Novel
PSEN1	c.1236+8 T>C	Splicing	0.33 – 1.76
PSEN1	c.A941G, p.E314G	Nonsynonymous SNV	0.91 – 2.5
PSEN2	c.A520G, p.M174V	Nonsynonymous SNV	0 – 0.02
PSEN2	c.T132A, p.T44T	Synonymous SNV	0.14 - 0.99
PSEN2	c.G185A, p.R62H	Nonsynonymous SNV	1.44 – 2.5
PSEN2	c.C423T, p.N141N	Synonymous SNV	0.41 – 0.66
PSEN2	c.C441T, p.S147S	Synonymous SNV	0.66 – 1.74
PSEN2	c.T708C, p.S236S	Synonymous SNV	1.83 – 3.39

PSEN2	c.G756C, p.A252A	Synonymous SNV	0.31 – 0.66
PSEN2	c.C771T, p.G257G	Synonymous SNV	0
PSEN2	C.G903T, p.T301T	Synonymous SNV	1 – 3.85

Only three individuals were found to harbour novel variants (not reported in 1000 genomes, EVS server or in >600 in-house control exomes) in these genes, all in *NCSTN*. These comprised one splice site substitution (*NCSTN* c.1352+1G>C), a frameshift deletion (*NCSTN* c.1010delA, p.Y337fs) and a nonsense mutation (*NCSTN* c.C349T, p.R117X). All three variants were confirmed by amplifying the relevant segments of DNA via PCR and performing standard Sanger sequencing (Figure 6.8).



**Figure 6.8.** Sanger sequencing validated the WES results and confirmed the presence of novel heterozygous variants in *NCSTN* in three individuals. *NCSTN* c.C349T, c.1010delA and c.1352+1 G>C (each mutation shown by an arrow in the respective figures) were verified by Sanger sequencing *NCSTN* exons 4, 9 and 11 respectively.

The NCSTN c.1352+1G>C splice site substitution is of the highly conserved first base of the *NCSTN* exon 10 donor splice site. This is predicted to be damaging (splice site abolished). The pathogenicity of this variant is further supported by a previous report of a different substitution at the same site (c.1352+1G>A) in a Chinese family with HS (full co-segregation demonstrated)<sup>137</sup>. The frameshift deletion and nonsense variants are both predicted to result in premature truncation (p.Y337fsX19, p.R117X) and are therefore likely to be pathogenic. The pathogenicity of the p.R117X variant is supported by the same mutation being reported in a Chinese HS kindred<sup>135</sup>. On the basis that these three variants are likely to be pathogenic, the relevant individuals were excluded from further exome sequencing analysis.

Of the remaining rare variants (<5% population frequency), none would be predicted to result in premature truncation. Two splice site variants were detected (*PSEN1* c.1236+8T>C, *APH1B* c.285-7A>T) with a population frequency of 0.3-1.8% and 0-0.3% respectively, neither of which were predicted to affect splicing. Ten nonsynonymous variants were detected, one of which (*APH1A* c.C735G (p.C245W)) was predicted to be potentially damaging (SIFT score 0, Polyphen not scored). *APH1A* c.C735G has a frequency of 0.37-0.61% in 1000 genomes, EVS and in house controls and it is hard to predict the pathogenic potential of this variant. On the basis of this, this individual was taken forward for further exome sequencing analysis. None of the 12 synonymous variants identified within the gamma-secretase genes fell within 10 bases of a splice site or were predicted to alter splicing. All individuals with synonymous variants were therefore taken forward for further exome sequencing analysis.

Exome sequencing only provides data relating to the coding regions of the genome and associated splice site boundaries. Sequences relating to regions such as promoters, enhancers, silencers and branch sites are not specifically targeted. Potentially pathogenic variants in those regions could not therefore be excluded in this cohort. All

easily identifiable variation in the gamma-secretase genes was however examined prior to interrogating the data to find novel disease-causing genes.

### **6.3.3 Familial analysis**

#### **6.3.3.1 Linkage analysis**

Selected affected and unaffected individuals from the six pedigrees were genotyped using the Illumina Infinium HumanExome 12v1.1 chip to generate linkage data (individuals shown in Figure 6.3). All regions of linkage are shown in Appendix 5. These areas represent regions in which the data is consistent with linkage (LOD score > -2). In contrast, linkage is excluded in the regions that are not listed. This data was generated to facilitate more efficient filtering of exome sequencing results by reducing the respective search space.

#### **6.3.3.2 Combined WES and linkage analysis of pedigrees 1-6**

Analyses were performed as stated in Figure 6.6. The results of analyses 1 and 2 are presented for each of the kindreds studied.

##### **6.3.3.2.1 Pedigree 1**

###### *6.3.3.2.1.1 Analysis 1*

No novel variants were shared between all four individuals subjected to exome sequencing. No filtered CNV was shared between all four individuals.

###### *6.3.3.2.1.2 Analysis 2*

Three variants with a population frequency of less than 1% were shared between all four individuals, two of which fell within identified regions of linkage (see Table 6.4).



### **6.3.3.2.2 Pedigree 2**

#### **6.3.3.2.2.1 Analysis 1**

Fifteen novel protein-altering or splice site variants were shared between the three members of this kindred who were exome sequenced, two falling within regions of linkage defined by genotyping (Table 6.4). No filtered CNV was shared between all three individuals.

#### **6.3.3.2.2.2 Analysis 2**

Five additional variants with a population frequency of < 1% were shared between those three individuals that fell within regions of linkage (Table 6.4).

### **6.3.3.2.3 Families 3-6**

The structure of the remaining pedigrees, combined with the available patient resources, meant that variants could not be filtered down to the minimal numbers seen in families 1 and 2 using the analysis 1 and 2 algorithms described in Figure 6.6. Only one individual was exome sequenced from Families 5 and 6 but all available family members were genotyped to determine regions of linkage within those families. The number of variants remaining following analyses 1 and 2 in pedigrees 3-6 is shown in Table 6.4.

**Table 6.4. A table showing the number of potentially disease-causing variants identified by analyses 1 and 2 in pedigrees 1-6.**

	Number of variants detected via analysis 1	Number of additional variants detected via analysis 2
Family 1	0	2
Family 2	2	5
Family 3	35	239
Family 4	12	29
Family 5	17	43
Family 6	34	119

### **6.3.3.3 Combinational analysis of data generated from all six pedigrees**

Upon analysing all of the novel variants and CNV that were identified within defined regions of linkage within each of the six kindreds there were no genes harbouring novel variants or affected by CNV that were shared between families. Upon performing the same analysis, but studying variants with a population frequency of < 1% there were three genes containing variants that were each shared between two kindreds.

### **6.3.3.4 Summary**

These familial analyses have highlighted a number of candidate variants and genes that may underlie the development of HS in these pedigrees. Bioinformatic interrogation of these data is ongoing. It was notable that there was no one gene in which all 6 pedigrees shared novel or rare variation, implying further genetic heterogeneity in HS.

#### **6.3.4 Analysis of phenotypically grouped, unrelated individuals with HS**

Analyses 1 and 2 were performed on data generated from individuals within the respective phenotypic sub-groups detailed in section 6.2.2. The hypotheses being tested were that a) all individuals will share novel protein altering variation, splice site variation or CNV affecting the disease-causing gene (analysis 1), b) all individuals will share rare (<1% population frequency) protein altering variation, splice site variation or CNV affecting the disease-causing gene (analysis 2). Neither hypothesis was proven in any of the sub-groups analysed. Within each group, the maximum number of individuals who shared variation in any one gene is summarised in the following sections.

#### 6.3.4.1 HS/arthritis cohort

##### 6.3.4.1.1 Analysis 1

**Table 6.5.** A table showing the distribution in the number of genes in which novel genetic variation was common to X number of individuals within the HS/ arthritis cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. Individuals with novel variation in gene: (Total no. in group = 7)</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
No. genes	4	5	78	1065

##### 6.3.4.1.2 Analysis 2

**Table 6.6.** A table showing the distribution in the number of genes in which rare (<1% population frequency) genetic variation was common to X number of individuals within the HS/ arthritis cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. individuals with rare variation in gene (Total no. in group = 7)</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
No. genes	4	23	54	248	1081	3755

### 6.3.4.2 Sinus tract disease cohort

#### 6.3.4.2.1 Analysis 1

**Table 6.7.** A table showing the distribution in the number of genes in which novel genetic variation was common to X number of individuals within the sinus tract disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. individuals with rare variation in gene (Total no. individuals = 30)</b>	<b>17</b>	<b>11</b>	<b>10</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>
No. genes	1	2	1	2	3	6	11	41	150

#### 6.3.4.2.2 Analysis 2

**Table 6.8.** A table showing the distribution in the number of genes in which rare genetic variation (<1% population frequency) was common to X number of individuals within the HS/ arthritis cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. individuals with rare variation in gene (Total no. individuals = 30)</b>	<b>23</b>	<b>19</b>	<b>17</b>	<b>16</b>	<b>15</b>	<b>14</b>	<b>13</b>	<b>12</b>	<b>11</b>	<b>10</b>	<b>9</b>	<b>8</b>	<b>7</b>
No. genes	1	1	3	1	3	6	5	6	13	19	32	41	90

### 6.3.4.3 Nodular disease (hormonal) cohort

#### 6.3.4.3.1 Analysis 1

**Table 6.9.** A table showing the distribution in the number of genes in which novel variation was common to X number of individuals within the nodular disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. individuals with novel variation in gene: (Total no. in group = 15)</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
No. genes	1	1	2	26	201	1960

#### 6.3.4.3.2 Analysis 2

**Table 6.10.** A table showing the distribution in the number of genes in which rare genetic variation (<1% population frequency) was common to X number of individuals within the nodular disease cohort (variation did not have to be exactly the same, it included any protein-altering, splice site or CNV affecting the relevant gene).

<b>No. individuals with rare variation in gene (Total no. individuals = 15)</b>	<b>11</b>	<b>10</b>	<b>9</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>
No. genes	1	1	1	3	9	24	30	111	339

### 6.3.4.4 Summary

There was no subgroup in which all individuals shared novel or rare variation in any one gene however genes in which variation was identified in multiple individual patients could be considered as potential disease-causing candidates. Bioinformatic analysis of these data is ongoing.

## **6.4 Discussion**

Work in chapters 3 and 5 implied that, in addition to the involvement of *NCSTN*, *PSENEN* and *PSEN1*, there may be further genetic heterogeneity underlying HS. Consequently, the aim of this study was to identify further disease-causing genes using a combination of positional cloning and next generation sequencing technologies. The patient cohorts being studied were integral to the potential success of this work, thus great care was taken in phenotyping both familial and unrelated cases. Two broad groups of individuals were recruited to the study, related individuals from multiplex kindreds with HS and unrelated individuals carefully grouped by phenotype.

### **6.4.1 Three novel mutations in *NCSTN* were identified through WES**

Of the 66 individuals studied, 30 had partaken in prior studies (Chapters 3 and 5) and had subsequently been screened for mutations in *NCSTN*, *PSENEN* and *PSEN1*. Of the remaining 36 individuals (all unrelated cases), three were found to harbour likely pathogenic variants in *NCSTN* by WES (no likely pathogenic variants were detected in *PSEN1* or *PSENEN*). One could therefore conclude, bearing in mind that 33 of those 36 cases reported a family history of HS (including all three individuals with *NCSTN* mutations), that around 9% (3/33) of British familial cases of HS arise as a consequence of mutations in *NCSTN*, *PSENEN* or *PSEN1*. This figure is slightly higher than that relating to the general disease population (<7%, presented in Chapter 5) which is perhaps not surprising given the familial nature of the group. These data require validation in a larger patient cohort.

### **6.4.2 Familial studies identify potential candidate variants**

This study demonstrates the enormous amount of data that is generated through WES and highlights the challenges faced in analysing and interpreting such information in the context of autosomal dominant inheritance and apparent genetic heterogeneity. Familial studies provide one way of circumnavigating genetic heterogeneity, as one would

expect only one disease gene in any single kindred. The success of such studies however depends upon the size, nature and availability of samples from individuals within the relevant pedigree. The structure of pedigrees 1 and 2 suggested that WES of multiple genetically distant individuals could significantly reduce the number of candidate disease-causing, shared variants. Data filtering (using the documented algorithms) further reduced the number of candidate variants to the extent that in pedigree 1 only three rare variants were shared between all four individuals subjected to WES (two within regions of linkage). The size and architecture of pedigree 2 meant that, despite multiple family members being sequenced, the list of potentially pathogenic variants could not be filtered down to the extent seen with pedigree 1. There were however only two novel variants and five rare variants shared between individuals. The nature and number of individuals sequenced from pedigrees 3-6 meant that variants could not be filtered down to the small numbers seen with pedigrees 1 and 2.

#### **6.4.3 Exclusion of genetic homogeneity outside of the gamma-secretase genes**

A combinational analysis of data generated from all six pedigrees revealed that there was no single gene in which all six pedigrees shared novel or rare variation. Different pairs of pedigrees shared rare variation in a total of three genes. These data suggest that there is no genetic homogeneity outside of the gamma-secretase genes (assuming pathogenic variants would be novel or rare), implying significant further genetic heterogeneity in familial HS.

#### **6.4.4 Analysis strategies employed to interpret WES data**

WES has revolutionised gene sequencing capabilities, yet the success of implementing this technique in the discovery of new disease genes is dependent upon the efficiency and accuracy of the bioinformatic analysis techniques that are available to analyse the vast amounts of data produced. At present, and dependent upon the nature, inheritance and genetic heterogeneity of the condition in question, bioinformatic approaches are



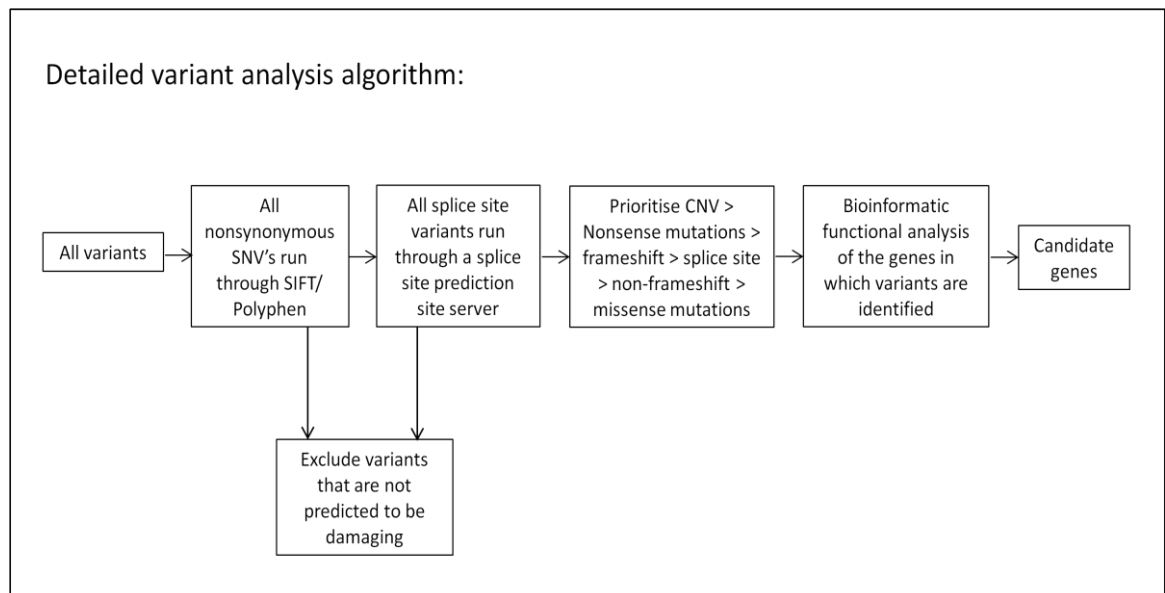
frequently insufficient to discern disease-causing mutations. This is partly because the efficiency of any such analyses depends upon the accuracy and completeness of genome databases, the resource against which all newly identified variants are compared to generate a measure of novelty. These databases are becoming increasingly comprehensive yet, at present, they largely comprise data derived from European Caucasian individuals, posing a significant challenge when analysing variation in individuals of other ethnicities (individuals of Afro-Caribbean and Asian ethnicity were included in this study). The completeness of data on these servers will improve with time and it is interesting to consider that the term “novel”, whilst useful in facilitating data filtering in this study, is likely to become increasingly redundant as even the rarest variants become documented in these servers.

Given the enormous quantity of data generated from the individuals studied in this project, analysis algorithms had to be designed with the primary aim of filtering out variants of unlikely pathogenicity. Those algorithms however were created in the knowledge that it would be possible to unintentionally filter out disease-causing variation in the process. For example, it was assumed that all cases were single gene disorders inherited in an autosomal dominant manner (thus homozygous variants were excluded), but that may not be the case. Furthermore, synonymous variants were excluded in the analysis 1 and 2 algorithms documented above. That was done in the knowledge that they are less likely to be pathogenic, not that they can never be pathogenic. Whilst generally considered as non-protein altering, it is well known that these variants can disrupt existing splice sites and create new splice sites, potentially conferring significant functional consequences<sup>287</sup>. Even accounting for the potential errors in this process, it is still critical to devise filtering algorithms that reduce the number of variants to a number that can practically be analysed in a more comprehensive manner. As pedigrees 3-6 and the large phenotypically defined, but

unrelated, cohorts demonstrated, even with the present algorithms in place there are still an enormous number of variants left to analyse.

#### **6.4.5 Potential analysis strategies that could be employed in the future interrogation of these data**

After data filtering (using algorithms stated in this chapter), the main mechanism by which candidate genes were identified in this study was by considering sharing of variation across either a pedigree or across a phenotypic subgroup of unrelated individuals. This has created variant lists which can now be further refined by adopting a more in-depth analytical approach. This may include assessing the nature of the variants (ie. nonsense variants are more likely to be more significant than missense variants etc.), conservation, predicted functional outcomes and known function of the relevant gene (potential algorithm of assessment shown in Figure 6.9). Whilst this kind of approach may be beneficial there are a few caveats. Firstly, using the above filtering criteria would further increase the risk of filtering out pathogenic variants. This is not just because yet more filtering steps are being added, but also because of the nature of those steps. Predictions regarding the functional consequences of mutations are only predictions, based on often limited knowledge about any given base and corresponding amino acid. Furthermore, whilst considering the function of the relevant genes may be helpful, it is important to remember that gamma-secretase gene variants would likely have been ruled out on that basis as the only known disease association at the time was FAD. The second problem with this overall kind of analytical approach is that, in the context of the groups studied here, it would still leave a large number of potentially pathogenic variants in the majority of analyses.



**Figure 6.9. A proposed detailed variant analysis algorithm.** This is a potential algorithm by which variants identified by WES and filtered using the algorithms described above could be further interrogated.

Another angle from which to approach data analysis is to acknowledge that some genes are more polymorphic than others. An analysis based purely on the sharing of different variants between individuals does not account for that variability. Some genes, whether because they are just larger or perhaps less functionally important, are more polymorphic. A more accurate approach would therefore be to acknowledge the general frequency of genetic variation in individual genes when prioritising candidate variants. One mechanism by which this analytical challenge can be addressed is currently being developed within the department. This in-house filtering programme will be able to compare variation in any given gene against variation in that gene in a control cohort (a gene-centric case control analysis). For example, if 50% of any given disease cohort (eg. sinus tract HS) had rare variation in gene X and this was unrelated to the clinical phenotype, then one would expect a similar frequency of variation in control cases. If *PSENEN* were the disease-causing gene in that group however, due to the conserved nature of that gene, the likelihood of individuals within the disease group sharing

variation in the gene versus the control group is much higher, thus identifying it as a candidate. Statistical probabilities can therefore be calculated on this basis, creating a new mechanism by which candidate genes can be analysed. The data generated in this study is currently being analysed in this way.

Another complexity in data analysis is highlighted by the identification of gamma-secretase mutations in HS. The above mechanisms of data analysis all assume that one gene is responsible for the disease in any given cohort. That was of course not the case with the gamma-secretase mutant group where three entirely independent genes were identified yet they were all functionally linked by their integration and function in the gamma-secretase complex. Thus, in future data analyses of apparently genetically heterogenous conditions it is important to identify whether variants are arising in groups of functionally related genes. One mechanism of integrating this into any analysis is by using protein-protein interaction software. This is capable of functionally grouping variants based on known protein functions, however much of the existing software is in its infancy and requires further development. One significant limitation with such programmes is that they are entirely dependent upon whether any research has been performed to link any two or more given proteins. It should never the less be considered, as without employing such a method, one would potentially need enormous numbers of unrelated individuals to detect three independent but functionally related genes. This problem was only circumnavigated in the initial identification of gamma-secretase genes in Chinese populations due to the number of pedigrees available for study and the structure and number of affected members within each kindred<sup>135</sup>.

When studying large groups of unrelated individuals, very accurate phenotyping is one further way to try and minimise genetic heterogeneity in any given study cohort. Assuring phenotypic homogeneity is difficult in the absence of a validated phenotypic classification however and, as was the case in this study, any attempt to clinically group

individuals cannot be evidence based. Furthermore, in a condition with such subtle phenotypic variation, segregating individuals on a purely clinical basis may be very difficult. One such example is that whilst the HS/ inflammatory arthritis cohort was interesting and consistent, there is nothing to say that the arthritis was not purely reactive and in no way related to the genetic basis of their HS. A further obstacle to accurate phenotyping is that, as again demonstrated by the gamma-secretase mutations, multiple functionally related genes can underlie the same phenotype (the phenotype of individuals harbouring gamma-secretase gene mutations is reviewed in Chapter 7 and put in to the context of ongoing clinical efforts to refine disease classification). Thus, even when studying a truly distinct clinical phenotype, there may still be a degree of genetic heterogeneity confounding the analysis. Ultimately, the number of unrelated individuals that are required to facilitate the identification of a disease-causing gene is partly dependent upon the amount of genetic heterogeneity within the cohort.

#### **6.4.6 Conclusion**

In summary, these data demonstrate that there is no single disease gene that is shared between pedigrees or shared between individuals within any of the distinct phenotypic subgroups. This implies that there is significant further genetic heterogeneity underlying HS. The data generated in this study is a valuable resource worthy of further bioinformatic interrogation and the algorithms devised in this project represent an important first step in data analysis. Ongoing statistical, network and pathway based analyses have the potential to further facilitate data filtering and combined with the mechanisms of in-depth analysis proposed in Figure 6.9 harbour the potential to identify new disease-causing genes in HS.

# **Chapter 7: Phenotypic analysis of individuals harbouring mutations in the gamma-secretase genes *NCSTN* and *PSENEN***

## **7.1 Introduction and aims**

Inter-individual variability in cutaneous features, disease distribution and associated systemic symptoms implies phenotypic heterogeneity in HS (discussed in section 1.1.4). Clinical efforts to categorise the disease remain at a premature stage but a recent study proposed three phenotypic sub-groups, “axillary-mammary”, “follicular” and “gluteal”<sup>30</sup>. Individuals with gamma-secretase mutations represent the first genetically defined subgroup of HS.

### **7.1.1 Aims and objectives**

The aim of this study was to present a case series of all mutation-positive individuals identified as part of this programme of research and to review the clinical phenotype of all such globally reported mutation-positive cases.

## 7.2 Patient cohort

A total of ten novel or rare (<1% population frequency) sequence altering gamma-secretase variants have been reported in this thesis, seven in *NCSTN*, one in *PSENEN*, one in *PSEN2* and one in *APH1A*. These comprise three missense, four splice site, two frameshift and one nonsense mutation. The missense changes in *NCSTN*, *PSEN2* and *APH1A* and two of the splice site substitutions in *NCSTN* are of uncertain pathogenicity (discussed in sections 5.5, 5.6 and 6.3.2). Individuals harbouring those variants were therefore excluded from this study on the basis that their clinical phenotype may not be related to the gamma-secretase genes. Two affected individuals within the same family were found to harbour the *PSENEN* mutation, both were taken forward for phenotypic analysis. A total of six individuals were therefore analysed, representing all of the likely pathogenic gamma-secretase variants that have been identified as part of this programme of research (3 females, 3 males, associated variants shown in Table 7.1).

**Table 7.1. All of the likely pathogenic gamma-secretase gene variants identified in this study.**

<b>Patient ID</b>	<b>Gene</b>	<b>Mutation</b>	<b>Nature of mutation</b>	<b>Predicted protein</b>
HS-01	<i>NCSTN</i>	c.C349T	Nonsense	p.Arg117X
HS-02	<i>NCSTN</i>	c.1010delA	Frameshift deletion	p.Tyr337Serfs*20
HS-03	<i>NCSTN</i>	c.1125+1 G>A	Splice site substitution	p.Glu333_Gln367del
HS-04	<i>NCSTN</i>	c.1352+1 G>C	Splice site substitution	Unknown
*HS-05a	<i>PSENEN</i>	c.66_67insG	Frameshift insertion	p.Phe23ValfsX98
*HS-05b	<i>PSENEN</i>	c.66_67insG	Frameshift insertion	p.Phe23ValfsX98

\*Denotes that the individuals are related (HS-05a is the daughter of HS-05b).



### **7.3 Clinical phenotype of patients harbouring gamma-secretase gene mutations**

**HS-01:** An Afro-Caribbean male who developed paired comedones, painful erythematous nodules, cysts, abscesses and sinus tracts in the post-auricular area, axillae, groin, buttocks and genitalia from the age of 32 (Figure 7.1, HS-01). He presented with features consistent with nodulocystic acne over his face, neck and trunk, had a BMI 27 and was an active smoker.

**HS-02:** A Caucasian male developed paired comedones, painful erythematous nodules, cysts, abscesses and scarring over the neck, post-auricular region, axillae, back, groin and buttocks from the age of 17. He also presented with an atypical, speckled, macular hyperpigmentation in axillae and groin (Figure 7.1, HS-02). He reported a history of acne vulgaris, had a BMI of 34 and was an active smoker.

**HS-03:** A Caucasian male developed paired comedones, pustules, cysts, erythematous nodules and abscesses over the neck, axillae, suprapubic area, groin, buttocks and medial thighs from the age of 16. He had features consistent with nodulocystic acne over his neck and back, his BMI was 30 and he was active smoker (no clinical photos available).

**HS-04:** A Caucasian female developed paired comedones, painful erythematous nodules and abscesses in the post-auricular region, groin and buttocks from the age of 20. She reported a history of acne vulgaris, had a BMI of 27 and was a non-smoker (no clinical photos available).

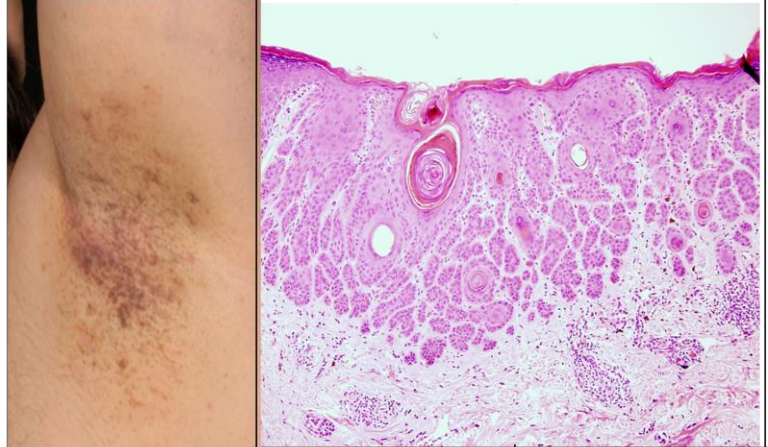
**HS-05a/05b:** HS05a is the daughter of HS05b. These Caucasian females developed paired comedones, pustules, painful erythematous nodules and abscesses in the axillae, groin and buttocks from the age of 15. HS-05a had additional sinus tract formation and medial thigh involvement and HS-05b had additional sub-mammary involvement. Both

also presented with an atypical, speckled, macular hyperpigmentation in the axillae and groin (Figure 7.1, HS05a, HS05b). HS-05a reported a history of acne vulgaris. They had BMI's of 23 and 33 and had never smoked.

HS-01  
(*NCSTN* p.R117X)



HS-02  
(*NCSTN* p.Tyr337Serfs\*20)



A) HS-05a  
B) HS-05a  
C) HS-05b  
(*PSENEN* p.Phe23ValfsX98)



**Figure 7.1. Clinical images of individuals with likely pathogenic gamma-secretase**

**mutations.** HS-01 had inflammatory nodules, sinus tracts and hypertrophic scars in the axillae and groin. HS-02 had atypical, speckled, macular hyperpigmentation in the axillae. Histopathological analysis of an axillary skin biopsy revealed hyperkeratosis, horn cysts and elongated thin filiform interconnecting epidermal strands extending into the superficial dermis. HS-05a has multiple suppurative sinus tracts and scars in both axillae. There was a faint speckled hyperpigmentation in both axillae which was more evident in the groin (no photo available). Her mother, HS-05b, had speckled hyperpigmentation, multiple comedones and one nodule in the left axilla. Photos courtesy of Guy's and St. Thomas' NHS Foundation Trust, all individuals consented to publication.

## 7.4 Discussion

This represents a small case series however a number of phenotypic features appeared common to the group. All six individuals presented with severe, persistent, widespread (involving “atypical” areas) and treatment resistant disease from a young age (average age 19, the average age of onset in the general disease population in 22.2<sup>9</sup>) and reported a family history of the condition. Common cutaneous features comprised paired comedones, painful erythematous nodules, abscesses and scars. Two of the more unique phenotypic features, and ones that may help in differentiating these individuals from the general disease population, included cystic acne-like changes in “atypical” areas (post-auricular area, neck and trunk) and an atypical speckled hyperpigmentation in the flexures. It should however be noted that those features were not universally present. A further noteworthy observation was that some individuals were non-smokers of normal BMI, which is unusual in the context of HS. Genetic and clinical details are summarised in Table 7.2.

“Typical” HS has been proposed to involve the axillae, sub-mammary region and groin whereas “atypical HS” comprises additional involvement of other sites<sup>30</sup>. All of these individuals would therefore be classified as having “atypical” disease. This terminology should be applied with caution however given that, as reported in Chapter 5, the gluteal region (an “atypical” site) was involved in up to 79% of cases studied<sup>22</sup>.

Flexural speckled hyperpigmentation was observed in three of the six individuals and arose prior to HS in all cases (distinguishing it from post-inflammatory pigmentation) and fully co-segregated with HS in the respective pedigrees. A fourth individual (HS-01) also had hyperpigmentation in the axillae and groin however the pigmentation was poorly defined and hard to characterise given the skin type. Cutaneous appearances clinically resembled those of Dowling-Degos Disease (DDD) which has been associated with mutations in *KRT5* and *POFUT1*<sup>71; 288</sup>. None of the three individuals with

pigmentary changes (HS02/05a/05b) harboured novel or rare (<1% population frequency) variants in *KRT5* or *POFUT1* on exome sequencing. DDD has frequently been reported to arise in association with HS yet, to our knowledge, no *KRT5* or *POFUT1* mutations have been reported in those cases<sup>70</sup> thus it is plausible that those individuals never had DDD but in fact harboured gamma-secretase gene mutations. Indeed, the clinical phenotype of those individuals very closely mirrors that of the mutation-positive subset reported here<sup>70; 289</sup>.

Analysis of all globally reported individuals with gamma-secretase mutations supports many of the phenotypic features observed in this group. All had severe and widespread disease. The majority had cystic involvement in the atypical sites mentioned above<sup>22; 109-113; 115; 116; 290</sup>. In total, including the mutations reported here, twelve of eighteen mutations reported in *NCSTN*, two of three mutations in *PSENEN* and the only *PSEN1* mutation appear to be associated with significant involvement of the neck and trunk<sup>22; 135-139; 141; 142</sup>. This would appear to be a prominent feature of the gamma-secretase phenotype. That said, there was no such involvement in three of the six patients studied here, thus it should not preclude one from considering more “typical” HS patients as harbouring gamma-secretase mutations. At least two of the globally reported individuals with *NCSTN* mutations appeared to have atypical pigmentary changes in the flexures, similar to that seen in some of our patients<sup>139; 142</sup>. It is notable that three individuals from globally reported mutation-positive multiplex kindreds have been reported to develop SCC in regions of HS<sup>135; 141</sup>. Furthermore, the father of HS03 developed SCC in a perianal region of HS and died from metastatic disease. This may be significant given that SCC is comparatively rare in the general disease population (5 of 2119 HS cases retrospectively analysed in one study)<sup>25</sup> and that heterozygous *NCSTN* knockout in mouse skin can induce the development of SCC<sup>291</sup>. It is not inconceivable therefore that

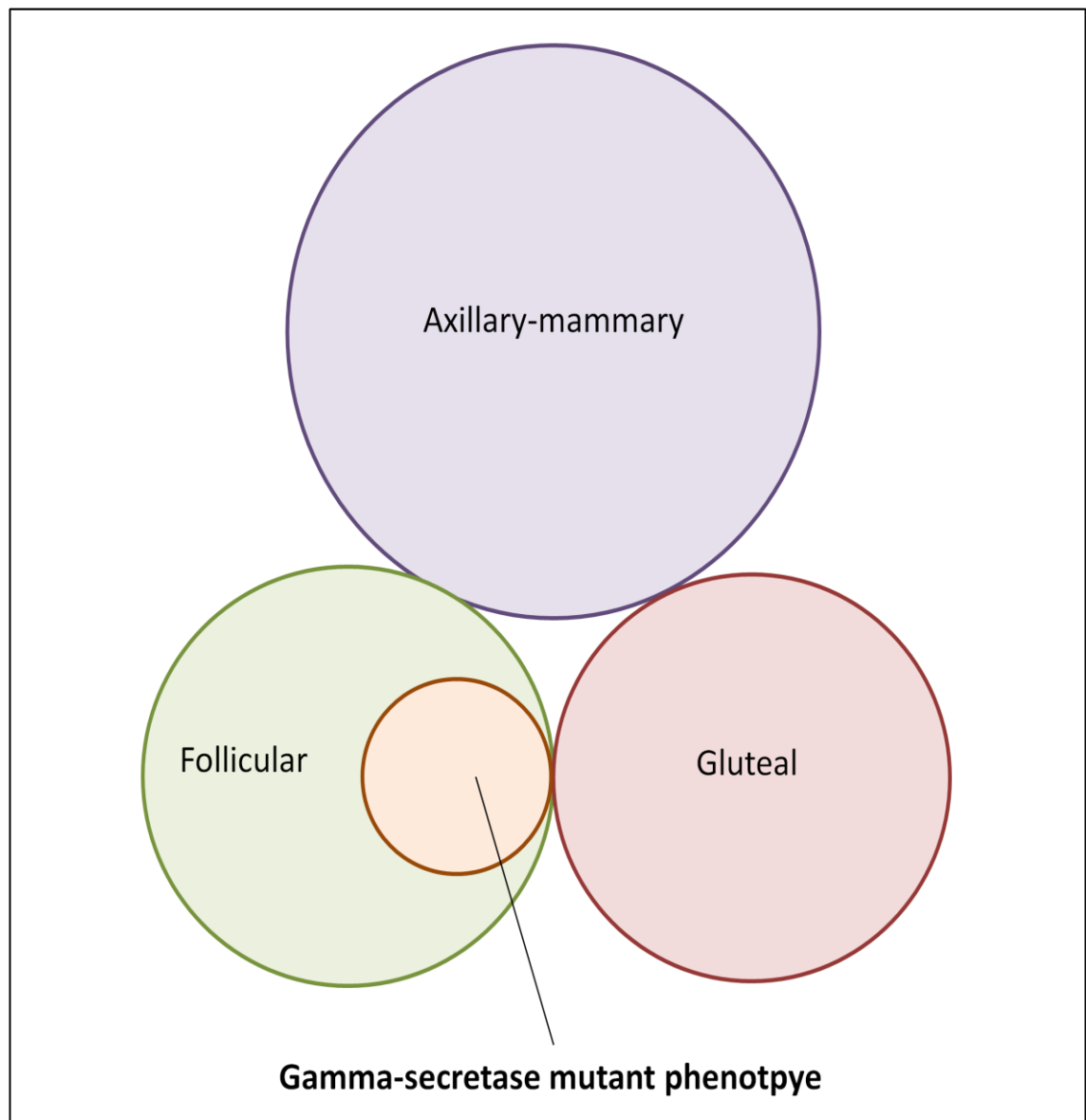
mutation-positive individuals may be at higher risk of developing SCC and this is a hypothesis worthy of future investigation.

A clinical approach to classification was recently adopted in France whereby they examined over 600 patients with HS and retrospectively grouped them according to defined clinical parameters. Latent class analysis identified three phenotypes; “axillary-mammary”, “follicular” and “gluteal”<sup>30</sup>. It is interesting that the individuals reported here appear to fall within the “follicular” group (an observation also reported by Ingram and Piguet<sup>292</sup>), partially validating that disease classification. Minor discrepancies in clinical characteristics (such as the percentage of patients with gluteal involvement and disease severity) and the proportion of total HS cases that fall within the gamma-secretase mutation and “follicular” groups (<7% vs. 26%) would however imply that there are likely to be multiple sub-phenotypes within the “follicular” cohort (Figure 7.2).

This represents a small case series and ongoing genotype-phenotype correlations are now required to refine the clinical characteristics associated with gamma-secretase gene mutations. We have now recruited over 300 cases for genetic analysis which will facilitate this process and potentially enable the future recognition and genetic testing of these individuals in the clinical setting. This may impact upon the stratification of patients for clinical studies, trials and hopefully, in the not too distant future, novel and more personalised treatments.

**Table 7.2. A table summarising the clinically distinct characteristics observed in individuals harbouring gamma-secretase gene mutations.**

Patient ID	Mutation	Early age of onset ( $\leq 20$ )	Family History	$\geq 6$ anatomical sites affected	“Atypical” sites affected (areas outside of axillae, sub-mammary region and groins)	Hurley stage II or III	Double-ended comedones, nodules abscesses and scars present	Flexural Pigmentation	Cutaneous features of nodulocystic acne
HS-01	c.C349T (p.Arg117X)	-	•	•	•	•	•	•	•
HS-02	c.1010delA (p.Tyr337Serfs*20)	•	•	•	•	•	•	•	•
HS-03	C.1125+1 G>A (p.Glu333_Gln367del)	•	•	•	•	•	•	-	•
HS-04	c.1352+1 G>C	•	•	•	•	•	•	-	-
HS-05a	c.66_67insG (p.Phe23ValfsX98)	•	•	•	•	•	•	•	-
HS-05b	c.66_67insG (p.Phe23ValfsX98)	•	•	•	•	•	•	•	-



**Figure 7.2** A diagram representing the different sub-types of HS that have now been described. Gamma-secretase mutation-positive patients appear to fall within the “follicular” phenotype as defined by Canoui-Poittrine *et al.*, 2013<sup>30</sup>.



## Chapter 8: Conclusions and future work

### 8.1 Introduction

HS is a debilitating and therapeutically challenging chronic inflammatory skin condition. The current paucity of effective treatment options partly reflects a poor understanding of disease pathogenesis. It can follow autosomal dominant inheritance in some families<sup>22; 39</sup>, thus genetic investigation represents one method by which to identify the molecular mechanisms driving disease development and propagation. The aim of this project was to investigate the genetic architecture of HS. In 2010, shortly after commencing this research, mutations were reported in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1* in six Chinese multiplex kindreds with HS<sup>135</sup>. This revelatory finding redefined the initial focus of this programme of research to investigating the potential role of gamma-secretase mutations in British pedigrees with HS and subsequently functionally characterising any mutations that were identified. The findings presented in this thesis confirm the involvement of gamma-secretase gene mutations in British familial cases of HS (Chapter 3), provide a novel insight in to the genetic mechanisms by which gamma-secretase mutations result in disease (Chapter 4), approximate the prevalence of gamma-secretase mutations in the general disease population (Chapter 5), contribute towards clarifying the gamma-secretase-associated HS phenotype (Chapter 7) and provide insight in to the complex and heterogenous genetic architecture underlying the condition (Chapter 6). Taken together, these findings have contributed towards characterising the first genetically defined sub-group of HS, conferring significant translational potential in individuals harbouring gamma-secretase mutations and potentially the wider disease cohort.

## 8.2 Gamma-secretase gene mutations in HS

Following the initial report of heterozygous gamma-secretase gene mutations in Chinese pedigrees with HS<sup>135</sup> these studies were the first to replicate those findings and to identify mutations in individuals of European descent (Chapter 3). A total of ten novel or rare (<1% population frequency) gamma-secretase gene variants have been identified as part of this project, 7 in *NCSTN*, 1 in *PSENEN*, 1 in *PSEN2* and 1 in *APH1A*. Five of these were predicted to be pathogenic (4 in *NCSTN*, 1 in *PSENEN*).

Globally, gamma-secretase gene mutations have now been reported in six Chinese, five British, one Japanese and three French multiplex kindreds as well as two apparently sporadic Chinese cases<sup>135-142</sup>. These comprise 18 mutations in *NCSTN*, three in *PSENEN* and one in *PSEN1*<sup>135-141</sup> of which five are nonsense mutations, nine result in frameshifts, five in altered splicing and three are missense mutations (Table 8.1). All but two of these mutations were reported in familial cases of HS and demonstrated complete penetrance in the pedigrees studied (two mutation positive cases without HS have been reported yet both were under the age of 12 and are thus too young to be confirmed as unaffected)<sup>135</sup>. No mutations have thus far been identified in *PSEN2*, *APH1A* or *APH1B* however screening has been limited<sup>141; 279</sup>.

**Table 8.1. All gamma-secretase mutations now reported in HS.**

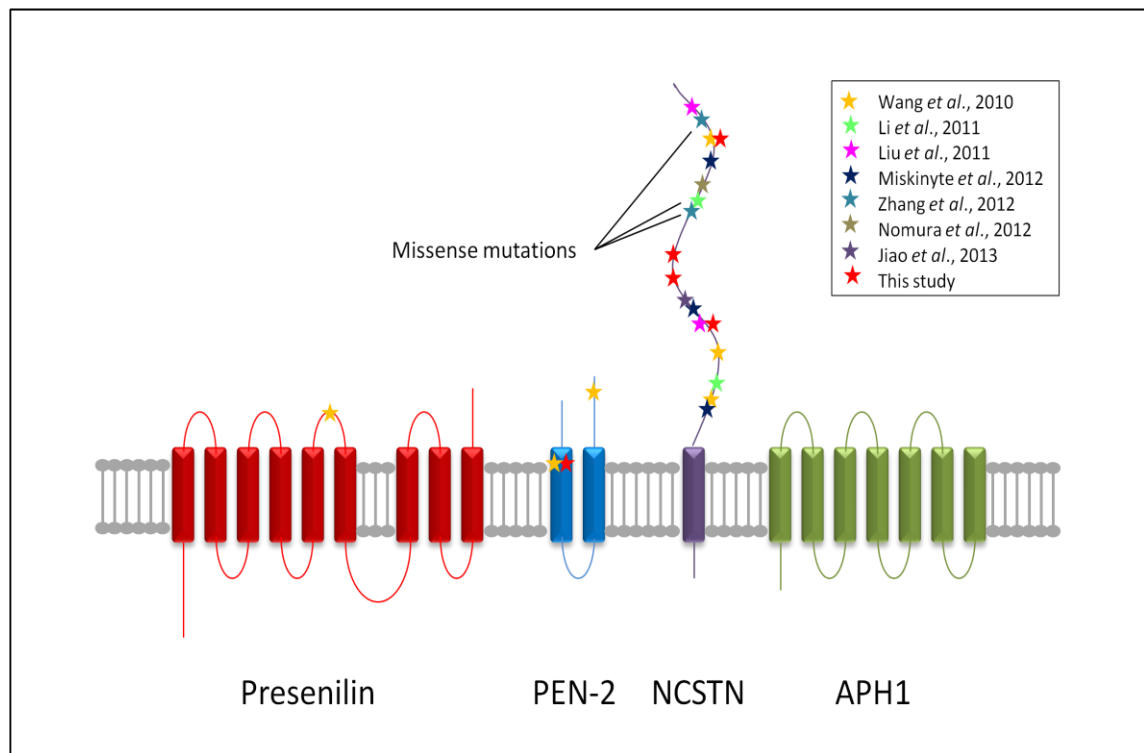
Gene	Mutation	Protein
<i>NCSTN</i>	c.210_211delAG <sup>4</sup>	p.T70fsX18
	c.223G>A <sup>5</sup>	p.V75I
	c.349 C>T <sup>1</sup>	p.R117X
	c.349C>T <sup>8</sup>	p.R117X
	c.487delC <sup>2</sup>	p.Q163SfsX39
	c.582+1delG <sup>6</sup>	
	c.632 C>G <sup>3</sup>	p.P211R
	c.647A>C <sup>5</sup>	p.Q216P
	c.1010delA <sup>8</sup>	p.Y337fs
	c.1125+1 G>A <sup>8</sup>	p.Q333_G367del
	c.1258C>T <sup>7</sup>	p.Q420X
	c.1300 C>T <sup>2</sup>	p.R434X
	c.1352+1 G>C <sup>8</sup>	
	c.1352+1 G>A <sup>4</sup>	
	c.1551+1 G>A <sup>1</sup>	
	c.1695 T>G <sup>3</sup>	p.Y565X
	c.1752delG <sup>1</sup>	p.E584DfsX44
	c.1768 A>G <sup>2</sup>	p.S590AfsX3
<i>PSENEN</i>	c.66delG <sup>1</sup>	p.F23LfsX46
	c.66_67insG <sup>8</sup>	p.F23VfsX98
	c.279delC <sup>1</sup>	p.F94SfsX51
<i>PSENI</i>	c.725delC <sup>1</sup>	p.P242LfsX11

1.Wang *et al.*, 2010 2. Miskinyte *et al.*, 2012 3. Li *et al.*, 2011 4.Lui *et al.*, 2011 5. Zhang *et al.*, 2012 6. Nomura *et al.*, 2012 7. Jiao *et al.* 2013 8. Identified as part of this project

### 8.2.1 A loss of function of the gamma-secretase genes appears to underlie HS

Functional analysis of the first two mutations reported in this thesis (*NCSTN* c.1125+1 G>C and *PSENEN* c.66\_67insG) revealed that they conferred very different effects on the nature of their respective mutant transcripts yet shared a downstream reduction in transcript abundance and corresponding protein expression. These data imply that haploinsufficiency is the common pathogenic mechanism by which both mutations are likely to result in HS. The nature of the globally reported mutations partially supports that mechanism as the majority are predicted to result in premature truncation which, in many cases, would be expected to induce NMD and thus reduce overall transcript abundance. Indeed, of those mutations in which RNA expression has been studied, many, but not all, were associated with a reduced transcript abundance<sup>135; 139</sup>. It should however be noted that the spectrum of global mutations and the associated functional data does not support a mechanism of haploinsufficiency in all cases, but it does potentially infer that loss of function of the gamma-secretase components is likely to be the unifying mechanism underlying disease development<sup>135-141</sup>.

It is noteworthy that the three missense mutations so far reported in HS (p.Val75Ile, p.Pro211Arg, p.Gln216Pro, )<sup>22; 138; 140</sup> are located within the *NCSTN* ectodomain. The site of these missense changes does not correspond with the highly conserved regions of the *NCSTN* ectodomain discussed in the introduction (amino acids 306-360, 419-458, 625-662) however mutagenesis of nearby amino acids (C213S and C230S) significantly affects gamma-secretase-induced APP cleavage and those amino acids are thought to play a key role in substrate specificity<sup>198</sup>. The exact mechanism by which these mutations result in disease is worthy of further exploration but the genetic location of the mutations supports a key functional role for the *NCSTN* ectodomain in the gamma-secretase complex and in HS pathogenesis.



**Figure 8.1. The location of all gamma-secretase gene mutations now reported in HS.** The majority of mutations have been reported in *NCSTN*. The three missense mutations correspond with amino acids within the ectodomain of *NCSTN*.

### 8.2.2 Cell membrane expression of the gamma-secretase components appears tightly regulated and gamma-secretase enzyme activity was not significantly impaired in mutant fibroblasts *in vitro*

The identification of mutations in genes encoding three of the four gamma-secretase complex components implicates complex function, rather than complex-independent roles of the individual components, in disease pathogenesis. To our knowledge, these are the first studies to investigate gamma-secretase protein expression and enzymatic function in HS and whilst this is preliminary data that needs to be replicated in a larger cohort of mutation-positive individuals, it provides a unique insight in to protein regulation and enzyme function in these individuals.

Immunoblotting data revealed that total cell protein levels of NCSTN and PEN-2 respectively appeared reduced in the two individuals harbouring mutations, yet that reduction was not replicated at the level of the cell membrane. These findings appear consistent with previous studies demonstrating that the production, maturation and export of the gamma-secretase components to the cell membrane is very tightly regulated<sup>189</sup>. There was subsequently no evidence to indicate a reduction in number or an impairment in maturity or activity of the gamma-secretase complexes present at the level of the cell membrane in mutant dermal fibroblasts. Indeed, if anything, raw endopeptidase enzyme activity per mg of solubilised membrane protein and carboxypeptidase processing of A $\beta$  peptides seemed very slightly increased in the individual harbouring the *PSENEN* mutation. Whilst this is an interesting observation, differences were slight and would require replication in that individual and other mutation-positive individuals for verification.

### **8.2.3 Immunohistochemical studies of human skin and *in vivo* gamma-secretase gene mutagenesis animal models support a role for gamma-secretase mutations in HS**

The anatomical distribution of NCSTN and PEN-2 in axillary skin mirrored sites of common histopathological changes in HS, thus supporting the involvement of gamma-secretase mutations in disease development. The widespread expression of gamma-secretase components however meant that few further conclusions could be drawn regarding the origins of HS development in the skin. Widespread expression was not unexpected given that the gamma-secretase complex is integral to such a wide range of essential cell processes.

The side effects experienced by individuals recently trialled on the gamma-secretase inhibitor Semagacestat (for Alzheimer's disease) also partially support gamma-secretase involvement in HS. Cutaneous side effects included non-specific skin rashes, hair

colour changes and an increased risk of skin cancer<sup>293; 294</sup>. It is unclear whether the skin changes resembled HS, however the alterations in hair colour would potentially support a biological role for gamma-secretase in the hair follicle.

*In vivo* animal models provide further evidence that the gamma-secretase components are involved in HS. Disruption of presenilin and nicastrin (*PSEN1* -/ *PSEN2* -, *PSEN1* - and *NCSTN* +/-) in the skin of mice results in follicular keratinisation, follicular atrophy, the formation of epidermal cysts, absent sebaceous glands and epidermal hyperplasia, all histological features of HS in humans<sup>291; 295; 296</sup>. Interestingly, these cutaneous manifestations initially appeared around the ears, eyelids, nose, mouth, inguinal and perineal areas. These, in part, correspond with cutaneous regions commonly affected in HS, namely the post-auricular, inguinal and perineal regions. Some nodules on the *PSEN1* - and *NCSTN* +/- mice underwent malignant transformation (developed in to SCC), which may correlate with the observation that multiple individuals with gamma-secretase gene mutations have developed SCC's in regions of affected skin. The apparent tumour-suppressor function of gamma-secretase in the skin is potentially mediated by its effects on Notch signalling (a known tumour suppressor) and the endocrine growth factor receptor (gamma-secretase inhibition results in an up regulation of EGFR which is up regulated in over 90% of head and neck SCC's)<sup>291</sup>.

A further interesting observation in these studies was that mice treated with a gamma-secretase enzyme inhibitor (LY-411,575), which leaves the gamma-secretase complex intact and at normal levels, develop similar lesions to NCT +/- mice<sup>291</sup>. This potentially implicates reduced enzyme activity, rather than a deficiency of any individual complex component, as a cause of the cutaneous manifestations observed in HS. This emphasises the need to further explore enzymatic function in mutation positive patients

#### **8.2.4 Gamma-secretase mutations may result in HS by affecting Notch signalling**

Taken together, the genetic evidence, the anatomical distribution of the gamma-secretase components in the skin and the phenotypic and histopathological similarities observed in mutagenised *in vivo* animal models firmly imply a role for the gamma-secretase components in disease pathogenesis. As well as further characterising the exact biological effects of gamma-secretase mutations on enzyme function it is also translationally important to investigate the downstream pathways through which these mutations result in HS. *In vivo* work in mice suggests that the skin changes observed in gamma-secretase deficient mice may be due to altered Notch signalling. The hair follicle and sebaceous glands of mice deficient in gamma-secretase are phenotypically identical to those of mice deficient in Notch 1 and 2 and the additional epidermal changes are present in mice deficient in Notch 1, 2 and 3<sup>295</sup>. Furthermore, and as eluded to above, gamma-secretase and Notch 1 deficiency have both been shown to induce SCC formation in mouse skin.

As discussed in the introduction, Notch is a trans-membrane receptor protein involved in cell to cell signalling. Gamma-secretase is integral to effective Notch signalling as, upon ligands binding to the extra-cellular domain of the Notch receptor, gamma-secretase cleaves the intracellular domain (NICD) which, in turn, re-locates to the nucleus and affects gene expression (Figure 1.9)<sup>297</sup>. Within the epidermis, Notch signalling is involved in stem cell clustering, epidermal cell differentiation (via the ligand delta-like 1), hair follicle differentiation (via jagged1) and epidermal proliferation<sup>298-300</sup>. It plays an integral role in maintaining hair follicles and sebaceous glands and acts as a tumour suppressor in the skin via its ability to control epidermal proliferation<sup>301</sup>. Notch signalling also appears to regulate inflammation via at least three mechanisms, 1) suppressing-TLR-4 induced cytokine expression in macrophages, 2) promoting IL-22 expression by CD4+ T cells and 3) inducing the production of NK



cells<sup>302-304</sup>. Consequently, any suppression of Notch signalling could induce or propagate the development of HS in two ways; 1) affect the structure and maintenance of the pilo-sebaceous unit and 2) alter cytokine profiles and contribute towards the inflammatory response observed in HS skin (IL-22 expression and NK production is reduced in HS<sup>90; 91</sup>). Notch signalling may therefore represent one downstream pathway through which gamma-secretase mutations, and any subsequent alteration in gamma-secretase function, could result in HS. It is potentially noteworthy that smoking has been shown to suppress Notch signalling in lung epithelium and if this were to confer any form of systemic effect then it may provide a hypothetical link between smoking and HS<sup>305; 306</sup>. It should, however, be noted that many of the individuals harbouring gamma-secretase mutations were non-smokers.

#### **8.2.5 Gamma-secretase mutations provide an insight in to the pathogenic mechanisms underlying HS**

In addition to implicating the gamma-secretase genes in disease development the discovery of gamma-secretase gene mutations may also inform matters of historical pathogenic importance such as the histological basis of the disease. As discussed in the introduction, HS was originally considered a disease of the apocrine gland yet recent histopathological studies have consistently highlighted follicular occlusion as the primary pathological phenomenon. The involvement of non-apocrine gland-bearing sites in HS, such as the sub-mammary folds, abdominal fold and medial thighs also partially supports those findings. The fact that gamma-secretase deficiency in mice, animals that do not have apocrine glands, results in follicular hyperkeratosis and cyst formation similar to that seen in human HS further implicates primary follicular involvement in HS, at least in the subset of patients harbouring mutations.

Primary immune dysregulation has recently been proposed to underlie HS<sup>82</sup>. There is little doubt, both clinically and histologically, that a significant inflammatory response

accompanies the disease yet it remains unclear whether that is a primary event, somehow resulting in the characteristic follicular occlusion noted on histology, or secondary to the build up of bacteria and keratin caused by initial follicular plugging. The identification of gamma-secretase mutations may inform this debate given that the follicular keratinisation and epidermal cyst formation observed in gamma-secretase deficient (*PSEN1* -/ *PSEN2* -) mice was rarely accompanied by an inflammatory response (<20% cases<sup>295</sup>). This suggests that follicular occlusion is likely to occur prior to any inflammatory response in individuals harbouring pathogenic gamma-secretase mutations.

#### **8.2.6 Gamma-secretase mutations underlie a minority of HS cases in the UK**

It is notable that the majority of familial and apparently sporadic cases of HS had no identifiable mutations or CNV in *NCSTN*, *PSENEN* or *PSEN1*. Only 2 of 7 multiplex kindreds, 3 of 32 familial cases and 3 of 48 cases sequentially recruited from a tertiary referral clinic harboured novel variants in these genes. Whilst this latter study may not be truly reflective of the general disease population, potentially over-representing severe disease, it is the largest study performed to date and provides an important estimate of mutational prevalence in HS. From these data we estimate that less than seven percent of British cases of HS are due to mutations in the gamma-secretase genes *NCSTN*, *PSENEN* and *PSEN1*. Given that there remains no evidence to implicate involvement of the other gamma-secretase genes, *APH1A*, *APH1B* or *PSEN2* in HS development, this figure probably reflects the prevalence of mutations in any of the gamma-secretase genes in HS.

### **8.2.7 A gamma-secretase mutation-related clinical phenotype is emerging which represents the first genetically defined phenotypic sub-group of HS**

In the knowledge that only a small proportion of HS cases appear to arise as a consequence of gamma-secretase gene mutations it was hypothesised that the mutation-positive group may harbour a specific and distinguishable clinical phenotype. Meticulous clinical examination of all individuals harbouring likely pathogenic gamma-secretase mutations combined with a review of the clinical data presented on all mutation-positive individuals in the literature, identified numerous phenotypic features common to the gamma-secretase mutant cohort. Some of these features, such as the atypical speckled flexural macular hyperpigmentation and superimposed cystic acne-like appearances in atypical areas, are relatively unique and have the potential to differentiate patients harbouring mutations in the clinical setting. It should be noted however that these phenotypic observations were limited by the small clinical sample size and that very few clinical features were consistent throughout the group. Indeed, even within the same pedigrees there was a degree of variation in phenotypic severity and cutaneous features. This phenotypic variation may, in part, be due to the many environmental and intrinsic factors that can impact on the condition, such as infection, smoking, obesity and hormonal factors. It is important that the clinical phenotype is continually refined as more mutation-positive cases emerge. This may ultimately improve the ability of clinicians to detect and genetically test these individuals in the clinical setting.

This work has contributed to initiating the long required process of classifying HS in to phenotypic sub-groups. In time, this may stimulate an alteration of the existing nomenclature to account for the variable pathogenic mechanisms underlying HS. This would raise clinical awareness that different sub-types exist, stimulate more disease-specific clinical studies and trials and ultimately potentially result in the application of

more effective differential treatments. That said, an over-arching terminology such as HS remains an important mechanism for grouping and recognising individuals such as this in the primary care setting, especially considering the already limited recognition of this disease in that environment.

### **8.2.8 There is no evidence to support an association between HS and FAD**

As discussed in the introduction (section 1.3.7), gamma-secretase mutations (*PSEN1* and *PSEN2*) were first identified in familial Alzheimer's disease (FAD). The vast majority of mutations in FAD are heterozygous missense changes but the exact mechanism by which they result in disease remains unclear. Both diseases appear to share a common genetic basis in as much as they affect genes involved in the gamma-secretase complex but it should be noted that only one mutation has been identified in HS in either of the presenilin genes implicated in FAD<sup>135</sup>. This mutation was in *PSEN1*, but mutations in this gene have not been replicated in this thesis or any other international study. No *PSEN2* mutations have been reported in HS to date. In contrast to the missense changes observed in FAD, the vast majority of mutations identified in HS are nonsense, frameshift and splice site mutations. The nature of the mutations identified in FAD (missense) might suggest that the mutant protein needs to be expressed in order to affect presenilin/ gamma-secretase activity<sup>294</sup> whereas, as we have demonstrated in this thesis, some of the mutations in HS appear to result in haploinsufficiency of the relevant genes. At an enzyme level, the unifying mechanism underlying FAD appears to be an alteration in carboxypeptidase-like gamma-secretase activity<sup>178</sup>. FAD mutations confer variable and often no effect on endopeptidase activity<sup>178</sup>. In HS, neither mutation studied in this thesis appeared to significantly affect endopeptidase activity and, in contrast to FAD, there was no statistically significant alteration in carboxypeptidase activity. Intriguingly, and in direct contrast with FAD, enzyme generated from *PSENEN* mutant fibroblasts appeared to generate reduced

A $\beta$ 42:A $\beta$ 40 ratios. This clearly requires replication in a larger cohort of mutation-positive individuals. A further consideration is that the FAD mutations appear to result in disease by affecting downstream APP processing, whereas in HS, the Notch signalling pathway potentially appears more relevant.

Aside from gamma-secretase enzymatic function, reports of mutations in only one complex component in FAD have lead some to propose that presenilin may act independently of the gamma-secretase complex to cause FAD. For example, *PSEN1* and *PSEN2* mutations can affect the expression and activity of  $\beta$ -secretase, another enzyme which is capable of converting amyloid precursor protein to  $\beta$ -amyloid<sup>307; 308</sup>, and can affect intracellular calcium levels<sup>191; 309</sup> (alterations in calcium fluctuations have been shown to occur prior to the development of amyloid plaques in AD)<sup>310</sup>. Mutations in multiple complex components would suggest that this is unlikely to be the case in HS.

The differing mechanisms that would appear to underlie these two disparate diseases go part way to explaining why there is no known clinical association between the two, or reports of FAD in the mutation positive multiplex kindreds with HS. The presence of HS and FAD in an individual or family should be interpreted with caution given the prevalence of both conditions in the population (prevalence of HS estimated to be up to 1%<sup>4</sup>). Even the detection of a mutation in *PSEN1* in such individuals would be inconclusive given the apparent heterogeneity in HS and one would need to establish firm co-segregation of FAD and HS within such families.

### **8.3 The future of gamma-secretase related research in HS**

The discovery of gamma-secretase gene mutations in HS suggests that the respective gamma-secretase complex components are involved in the pathogenesis of at least some cases of HS. Further functional characterisation of these mutations is now required to determine the specific genetic and molecular mechanisms underlying disease development and to realise the translational potential of these findings.

#### **8.3.1 Future functional characterisation of the gamma-secretase mutations identified in HS**

Two of the mutations characterised as part of this thesis appeared to result in haploinsufficiency, however, as previously mentioned, that would not appear to be the pathogenic mechanism underlying all globally reported mutation-positive cases. Whilst the spectrum of mutations might imply that loss of function is the common pathogenic mechanism, that remains to be proven. Careful elicitation of the exact mechanisms underlying all identified mutations, including the splice and missense variants, is required to verify this hypothesis.

A further hypothesis is that these mutations result in HS by affecting gamma-secretase complex function rather than through independent, and as yet potentially unidentified, mechanisms associated with the individual components. Whilst the endopeptidase and carboxypeptidase-like activities of the gamma-secretase complex represent two of the better understood and most extensively studied facets of the enzyme, it should not be forgotten that these represent but two of an increasing number of functions attributed to the complex. Gamma-secretase also appears to play a role in endocytosis, protein trafficking and phagocytosis and complex dysfunction may affect any of those processes. The *in vivo* data potentially implicating Notch signalling in HS would however imply that the endopeptidase or carboxypeptidase properties of gamma-secretase are indeed relevant in disease pathogenesis.

It is now important to assess the impact of a larger number and wider spectrum of gamma-secretase mutations on endopeptidase and carboxypeptidase enzyme activity. Many factors need to be considered in designing such experiments given the cost, potential time input, technical difficulty and sheer number of cells required for such work. Primarily one would need to decide upon the substrate or substrates to be studied. If it is hypothesised that Notch signalling is the relevant downstream pathway in HS then Notch, rather than APP, may be the optimal substrate to study, particularly given that some FAD mutations have been reported to affect substrate specificity<sup>311</sup>. Whilst Notch has been used in gamma-secretase activity assays, the relative inexperience with that substrate combined with a poor understanding of N $\beta$  processing (reflecting carboxypeptidase activity), means that data interpretation could be a significant challenge.

The *in vitro* APP based gamma-secretase activity assay described in this study is a standard, accepted and commonly used method of analysing enzymatic function but it represents just one of many potential methods to explore complex activity. A faster but slightly more indirect method of examining enzymatic function is to assess the expression of genes involved in signalling cascades downstream of gamma-secretase. One could look at the expression of single gene transcripts, eg. *HES1* (expression driven specifically by Notch signalling)<sup>272</sup> or the expression of multiple genes from a variety of signalling pathways, eg. via microarrays. This type of work would not only provide an insight in to enzymatic function but also contribute towards confirming the specific downstream pathways relevant in HS.

A further consideration when performing any of the above experiments is the cell type to be studied. Primary dermal fibroblasts were the source of gamma-secretase complexes in all enzyme activity assays reported as part of this thesis but immunohistochemical analysis of axillary skin revealed keratinocytes to be another cell

type worthy of future study (strong expression of NCTSN and PEN-2). The relative instability of primary keratinocytes in culture would however require careful consideration in this context. It may be beneficial to perform the above suggested experiments in a range of different cell types to establish any cell-type specific variation. In considering the source of these respective cells, whilst primary mutant cells represent a very valuable resource, preliminary experiments could initially focus on the effects of knocking down NCSTN or PSENEN (using SiRNA technology) in control cells (either primary or immortalised) on gamma-secretase enzyme activity and downstream signalling.

One of the intriguing and potentially revealing features of HS is the unusual cutaneous distribution. Gamma-secretase is widely expressed in the skin and internal organs and is integral to multiple essential cellular signalling pathways yet the heterozygous mutations reported in HS are only associated with a cutaneous disease with a very specific anatomical distribution. If these mutations significantly impaired or inhibited enzyme function then one would perhaps expect more severe and widespread consequences, corroborated by *PSEN1* knockout in mice resulting in mortality in embryo. The heterozygous nature of the mutations in HS may be a key factor in the disease distribution. Heterozygous knockout of PSEN1 (*PSEN1* -/+) in mouse skin confers no apparent effect<sup>296</sup>, yet full knockout (*PSEN1* -/-) results in the aforementioned skin changes<sup>285; 289; 290</sup>. This *in vivo* data implies that a heterozygous mutation alone, at least in that gene, animal model and particular environment, is not sufficient to induce disease. These data may relate to the results documented in this thesis stating that an apparent total cellular reduction in *NCSTN* or *PEN-2* expression was not sufficient to affect corresponding protein expression or enzyme function at the level of the cell membrane. One might therefore hypothesise that other exogenous or endogenous factors are required to induce disease. The full penetrance observed in all



globally reported kindreds implies that any such triggering factors are likely to be universal to everyone (eg. hormonal influences, apocrine secretions) rather than specific to individuals (eg. smoking). This is perhaps consistent with some of the mutation-positive patients described in this study being non-smokers of normal BMI, which is unusual in the context of HS. Accounting for the anatomical sites of HS these triggering factors may include heat, sweat, friction, apocrine gland or eccrine gland secretions and the local microbiome for example. Assuming that an alteration in enzyme function underlies disease development then one might imagine that one or a combination of these factors is sufficient to alter complex formation, maturation, turnover and ultimately function beyond a critical threshold required for normal follicular function. This would potentially explain why the entire skin surface is not affected in these individuals. It is noteworthy that pH, temperature and salinity have all been demonstrated to confer a significant impact on gamma-secretase enzyme function *in vitro*<sup>266</sup>. The flexural areas associated with HS are sites of skin occlusion (previously shown to significantly increase pH), sweat (potentially tying in the effect of salinity on enzyme function) and warmth<sup>122</sup>. It would therefore seem highly pertinent to investigate for any differential effects of temperature, pH and salinity on enzyme function in mutant versus wild type cells.

It may also be interesting to explore the extent and distribution of follicular occlusion in patients harbouring mutations. One might hypothesise that there is generalised follicular occlusion which only generates an inflammatory response in the presence of one or more of the factors mentioned above (eg. a specific microbiome in flexural areas). A further hypothesis worthy of exploration is that gamma-secretase complex composition may play a role in the unusual disease distribution. As discussed in the introduction, there are at least 6 different possible complexes and their differential location and distribution within the skin is as yet unknown. As the molecular mechanisms driving

disease development emerge in these cases it will be important to consider how gamma-secretase mutations and any potential subsequent alteration in gamma-secretase complex function relate to the immunological and inflammatory responses described in HS. Methods to explore this may include studying the effects of gamma-secretase mutations, or at least knockdown of the relevant constituents, on cytokine expression and dendritic cell differentiation, migration and function (given the emerging role of the hair follicle in facilitating the migration of Langerhan cell precursors in to the epidermis, discussed in section 1.2.1).

This recent genetic breakthrough and the contributory data presented as part of this thesis clearly generate many new and interesting hypotheses worthy of further investigation. Future studies need to harness these findings and make them translationally relevant by identifying new therapeutic targets. It should be born in mind when interpreting future functional studies that numerous therapies have already been developed to target the gamma-secretase/ Notch signalling axis including gamma-secretase inhibitors and modifiers.

### **8.3.2 The translational relevance of the above functional studies is partially dependent upon the prevalence of gamma-secretase gene mutations in the general disease population**

The relevance and impact of the above proposed studies depends somewhat on the proportion of total HS cases that are in any way related to the gamma-secretase components. Data generated as part of this thesis implies that gamma-secretase gene mutations are only present in a minority of HS cases, including cases from kindreds exhibiting autosomal dominant inheritance suggestive of single gene inheritance. That does not exclude the involvement of gamma-secretase, or at least gamma-secretase related pathways, in the remaining cases however. Some individuals may harbour as yet unidentified variation in the gamma-secretase genes or mutations in other genes that are

associated with the complex or encode proteins involved in up or down-stream signalling pathways. Two broad mechanisms by which this could be addressed would be to 1/ identify further disease causing genes and further clarify the genetic architecture of the condition, 2/ perform functional assays to assess downstream products produced as a consequence of gamma-secretase related pathways, as briefly mentioned above. Not only would such work provide a perspective on the relevance of ongoing translational work related to the gamma-secretase gene mutations but it could identify the specific molecular pathway or pathways through which the gamma-secretase mutations result in HS (potentially the Notch signalling pathway, as discussed above). Further genetic studies may alternatively reveal genes that are functionally independent of gamma-secretase. The first of those methods (attempting to identify new disease genes) was employed as part of this thesis. Next generation sequencing identified a number of candidate genes worthy of further exploration (described in Chapter 6). These results and the experimental methods that could be employed to further interrogate this genetically heterogeneous condition are discussed in Section 8.4.

### **8.3.3 Functional characterisation of the gamma-secretase gene mutations identified in HS may provide a novel insight into general gamma-secretase enzyme complex function**

As well as providing a basis upon which to further our understanding of HS disease pathogenesis, the gamma-secretase mutations identified in HS now represent a novel model through which to gain a better understanding of general gamma-secretase function. This knowledge could be of significant relevance in far wider reaches of medicine, including Alzheimer's disease. Deciphering the subtle mechanistic differences associated with HS and FAD mutations may provide a unique insight in to gamma-secretase complex formation, maturation, function and the mechanisms involved in processing substrates such as Notch and APP. Furthermore, affected regions

of skin are easily accessible in comparison with brain tissue in the context of FAD. A better understanding of complex biology is of broad physiological relevance. In the skin alone it may provide insight into basic hair follicle biology and the pathophysiology of related cutaneous conditions such as acne conglobate, acne vulgaris, dissecting cellulitis and pilonidal abscesses.

#### **8.3.4 Further investigations are required to determine if there is any association between HS and FAD**

As previously discussed, there is no clear association between the gamma-secretase gene mutations identified in HS and the development of FAD. That does not however exclude a functional relationship between the two conditions. If the mutations identified in HS are found to impair enzymatic endopeptidase activity or indeed alter carboxpeptidase activity such that the A $\beta$ 42:A $\beta$ 40 ratio is decreased then it could be hypothesised that they may in fact confer a protective affect against developing Alzheimers disease. Prospective clinical studies to address such a hypothesis would require very large numbers of patients. A pilot neuro-imaging study however, in which elderly mutation-positive individuals from affected HS kindreds and age matched healthy volunteers could be scanned for cranial amyloid deposition, would be one way to gather some preliminary data.

#### **8.3.5 Refinement of the gamma-secretase related clinical phenotype**

The phenotypic observations documented in Chapter 7 are based on a small cohort of individuals. The phenotypic characteristics need to be verified and statistically differentiated from the remaining HS population by studying a much larger group of patients. We have now recruited over 300 individuals with HS for ongoing genetic analysis and are collecting detailed phenotypic information on all patients attending our tertiary referral clinic. This should provide a firm basis from which to perform such work as well as facilitating ongoing phenotypic dissection of the remaining HS cohort.

As mentioned in chapter 7, it is notable that multiple individuals reported to harbour gamma-secretase mutations have developed SCC in regions affected by HS<sup>128; 134</sup>. SCC is comparatively rare in the general disease population<sup>25</sup>, although it should be noted that few HS patients have been retrospectively analysed and reported in the detail that these mutation-positive individuals have over the last three years. *In vivo* data implying that heterozygous *NCSTN* knockout in mouse skin can induce the development of SCC<sup>291</sup> would however potentially support a hypothesis that individuals harbouring gamma-secretase gene mutations are more prone to developing SCC in regions affected by HS than mutation-negative affected individuals. A prospective analysis of a large number of closely matched mutation-positive patients, mutation-negative patients and healthy control subjects would be the optimal way to address this hypothesis.

## **8.4 Ongoing genetic investigation of HS**

### **8.4.1 Further genetic heterogeneity in HS**

As discussed in Chapters 3 and 5, only a small proportion of HS patients appear to harbour mutations in the gamma-secretase genes. There are a number of potential explanations for these results. It may be that mutation-negative individuals harbour non-coding variants (for example in promoter, enhancer or silencer regions), undetected coding variants or undetected CNV affecting the gamma-secretase genes. This would seem unlikely to explain over 93% of disease cases however. An alternative explanation is that there is further genetic heterogeneity underlying HS. This hypothesis is supported by co-segregation data in two of the largest mutation-negative kindreds being inconsistent with involvement of *NCSTN*, *PSENEN* or *PSEN1* and also by the observation that there appear to be subtle phenotypic differences in the gamma-secretase mutation-positive cohort. Consequently, mutation-negative families and meticulously phenotyped clinical subgroups were studied using a combination of traditional and next generation gene mapping techniques with the aim of identifying further disease causing mutations and genes (Chapter 6).

### **8.4.2 A combination of traditional and next generation methods of genetic interrogation identified further potential candidate genes in HS and excluded genetic homogeneity outside of the gamma-secretase genes**

This study exemplified the quantity of data generated by exome sequencing and the challenges faced in analysing such data. Genetic interrogation of the two larger pedigrees highlighted a handful of potentially causative variants, which are a focus of ongoing study. The employed filtering strategies were unable to filter variants down to such small numbers in the remaining pedigrees and the phenotypically defined groups of unrelated individuals, emphasising the bioinformic challenges faced in analysing a genetically heterogeneous condition such as HS. It was noteworthy that there was no

single gene in which variation was shared between all familial cases, excluding genetic homogeneity outside of the gamma-secretase genes.

### **8.4.3 The future of genetic studies in HS**

Given the likelihood of further genetic heterogeneity in HS, genetic studies remain an important method by which to investigate disease pathogenesis. The genetic basis of a minority of cases is now known, yet the genetic architecture of the remaining HS population is undetermined. Up to 42% of HS cases report a family history of the condition but the influence of genetic susceptibility in the remaining cohort is unclear. The validity of clinically determined measures of heritability in HS, such as people reporting a family history of the condition, should be interpreted with caution. Factors such as under-reporting within families (potentially as a consequence of embarrassment or lack of family communication) and variability in disease inheritance or penetrance are potential confounders. From the data presented in Chapter 5 it may therefore be more appropriate to conclude that at least 42% of HS cases are familial. The proportion of HS cases that follow single gene mendelian inheritance, more complex polygenic inheritance and no genetic predisposition remains to be elicited.

The ever increasing armoury of genetic methodologies and bioinformatic approaches to data analysis will help to identify new disease genes in HS and facilitate a better understanding of the genetic architecture of the condition. Efforts to analyse the valuable exome sequencing data generated as part of this research are ongoing, comprising a more statistical (gene-centric case-control analysis) and network/ pathway (using protein-protein interaction software) based approach (discussed in Chapter 6). The knowledge that gamma-secretase mutations are involved in HS highlights a range of functionally related candidate genes to look out for in all ongoing analyses. That does of course assume that any further disease-causing genes encode proteins that are in some way related to that enzyme complex. Proteins integral to gamma-secretase

structure, assembly and activity are discussed in the introduction along with all currently known interactors and substrates. Genes encoding proteins integral to the complex such as *APH1A*, *APH1B* and *PSEN2* are clear candidates, however no mutations have been identified in those genes to date. This is partly consistent with functional work in mice demonstrating that *PSEN2* deficiency does not alter the ability of the gamma-secretase complex to cleave Notch and APP (in contrast to *PSEN1* deficiency)<sup>312</sup> and *PSEN2* and *APH1B* knockout mice have a normal phenotype but for the development of lung haemorrhages and fibrosis with age in *PSEN2* knockouts<sup>160; 212</sup>. *APH1A* appears more important given that *APH1A* knockout is lethal<sup>213</sup> and that recent functional work has demonstrated that *PSEN1*, *PEN-2* and *APH-1A* are vital for normal enzyme function<sup>273</sup>. *APH1A* would therefore seem the most likely of these genes to be involved in HS. Further genes to consider would be those involved in the Notch signalling pathway. Proteins involved in the core Notch signalling cascade are documented in the introduction. Protein interaction analyses mentioned above and in Chapter 6 provide a key mechanism by which exome data can be assessed in relation to genes within known candidate pathways such as the gamma-secretase/ Notch signalling axis.

Despite implementing comprehensive analysis algorithms and devising new data analysis strategies (statistical and protein/ pathway analyses) the degree of locus heterogeneity apparent in HS poses a significant challenge when attempting to identify new disease genes. A more familial approach, restricting the amount of genetic heterogeneity in any given sample, probably represents the most efficient method by which to identify new disease genes. A multidisciplinary approach to refining the phenotypic classification in HS would aid that process by facilitating the study and joint analysis of multiple but phenotypically similar pedigrees. This would also be integral to efficiently analysing any large unrelated patient cohorts. Clinical observations,



statistical group analyses, histo-pathological discriminators or immunological signatures have the potential to contribute towards defining these particular phenotypic sub-groups. As described in Chapter 7, a recent clinical study proposed a preliminary phenotypic classification in HS, dividing the disease in to three phenotypic groups. The gamma-secretase mutation-associated phenotype described here appeared to fall within the "follicular" group but indicated that there are likely to be further sub-phenotypes within that "follicular" cohort. Given the higher proportion of familial cases reported in that group, prioritising genetic investigation of the gamma-secretase mutation negative individuals harbouring the "follicular" phenotype may prove worthwhile. Upon deciding which phenotypic groups should be studied, patients could be further filtered based on factors implicating a firm genetic basis to their disease. These may comprise a clear autosomal dominant family history, early age of onset, non-smokers and individuals of normal BMI.

## **8.5 Concluding remarks**

This thesis supports the involvement of gamma-secretase gene mutations in familial HS and verifies that some familial cases of HS are single gene disorders demonstrating autosomal dominant Mendelian inheritance. It has demonstrated that haploinsufficiency is at least one pathogenic mechanism by which these mutations are likely to result in disease. The involvement of genes encoding three of the four gamma-secretase components implicates complex function in disease pathogenesis yet these studies revealed no significant alteration in gamma-secretase complex maturity or activity at the level of the cell membrane in mutation-positive individuals. The cohort of individuals with gamma-secretase mutations is the first genetically and increasingly phenotypically defined subgroup of HS but only appears to represent a minority of overall HS cases. There would appear to be significant further genetic heterogeneity underlying the condition and investigations are ongoing to identify more disease-causing genes. The nature of any such genes may highlight the up or downstream pathways through which the gamma-secretase complex is involved in HS, or may highlight entirely novel and independent molecular pathways. Whilst continuing phenotypic dissection of this clinically heterogeneous condition will potentially aid such studies it may be that subtly different clinical phenotypes will only be characterised as our genetic knowledge evolves. The data presented in this thesis provides a firm basis for ongoing functional investigation and highlights the gamma-secretase complex and potentially the gamma-secretase-Notch pathway as a potential therapeutic target in at least some cases of HS.

## References

1. Velpeau, A. (1839). Bechet Jeune Z: Dictionnaire de médecine, on Repertoire Générale des Sciences Medicales sous le Rapport Theorique et Pratique.
2. Verneuil, A.S. (1854). Etudes sur les tumeurs de la peau et quelques maladies de glandes sudoripares. Arch Gen Med 94, 693-705.
3. Schiefferdecker, B. (1922). Die Hautdrüsen der Menschen und der Säugetiere, ihre histologische und rassenanatomische Bedeutung sowie die muscularis sexualis.(Stuttgart).
4. Brunsting, H.A. (1939). Hidradenitis suppurativa; abscess of the apocrine sweat glands. A study of the clinical and pathologic features, with a report of twenty-two cases and a review of the literature. Arch Derm Syphilol 39, 108-120.
5. Shelley, W.B., and Cahn, M.M. (1955). The pathogenesis of hidradenitis suppurativa in man; experimental and histologic observations. A M A archives of dermatology 72, 562-565.
6. Yu, C.C., and Cook, M.G. (1990). Hidradenitis suppurativa: a disease of follicular epithelium, rather than apocrine glands. The British journal of dermatology 122, 763-769.
7. Boer, J., and Weltevreden, E.F. (1996). Hidradenitis suppurativa or acne inversa. A clinicopathological study of early lesions. The British journal of dermatology 135, 721-725.
8. Plewig, G., and Steger, M. (1989). Acne inversa (alias acne triad, acne tetrad or hidradenitis suppurativa) In Acne and related disorders, R. Marks and G. Plewig, eds. (London, Martin Dunitz), pp 345–357.
9. Von der Werth, J.M., and Williams, H.C. (2000). The natural history of hidradenitis suppurativa. Journal of the European Academy of Dermatology and Venereology 14, 389-392.
10. Nazary, M., der Zee, H.V., Prens, E., Folkerts, G., and Boer, J. (2011). Pathogenesis and pharmacotherapy of Hidradenitis suppurativa. European journal of pharmacology 672, 1-8.
11. Revuz, J.E., Canoui-Poitaine, F., Wolkenstein, P., Viallette, C., Gabison, G., Pouget, F., Poli, F., Faye, O., Roujeau, J.C., Bonnelye, G., et al. (2008). Prevalence and factors associated with hidradenitis suppurativa: results from two case-control studies. Journal of the American Academy of Dermatology 59, 596-601.
12. Jemec, G.B., Heidenheim, M., and Nielsen, N.H. (1996). The prevalence of hidradenitis suppurativa and its potential precursor lesions. Journal of the American Academy of Dermatology 35, 191-194.

13. Cosmatos, I., Matcho, A., Weinstein, R., Montgomery, M.O., and Stang, P. (2013). Analysis of patient claims data to determine the prevalence of hidradenitis suppurativa in the United States. *Journal of the American Academy of Dermatology* 68, 412-419.
14. Lookingbill, D.P. (1988). Yield from a complete skin examination. Findings in 1157 new dermatology patients. *Journal of the American Academy of Dermatology* 18, 31-37.
15. Vazquez, B.G., Alikhan, A., Weaver, A.L., Wetter, D.A., and Davis, M.D. (2013). Incidence of hidradenitis suppurativa and associated factors: a population-based study of Olmsted County, Minnesota. *The Journal of investigative dermatology* 133, 97-103.
16. Sartorius, K., Emtestam, L., Jemec, G.B., and Lapins, J. (2009). Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *The British journal of dermatology* 161, 831-839.
17. Harrison, B.J., Read, G.F., and Hughes, L.E. (1988). Endocrine basis for the clinical presentation of hidradenitis suppurativa. *The British journal of surgery* 75, 972-975.
18. Rompel, R., and Petres, J. (2000). Long-term results of wide surgical excision in 106 patients with hidradenitis suppurativa. *Dermatologic surgery : official publication for American Society for Dermatologic Surgery [et al]* 26, 638-643.
19. Konig, A., Lehmann, C., Rompel, R., and Happle, R. (1999). Cigarette smoking as a triggering factor of hidradenitis suppurativa. *Dermatology* 198, 261-264.
20. Simonart, T. (2010). Hidradenitis suppurativa and smoking. *Journal of the American Academy of Dermatology* 62, 149-150.
21. Jemec, G.B. (2012). Clinical practice. Hidradenitis suppurativa. *The New England journal of medicine* 366, 158-164.
22. Pink, A.E., Simpson, M.A., Desai, N., Dafou, D., Hills, A., Mortimer, P., Smith, C.H., Trembath, R.C., and Barker, J.N. (2012). Mutations in the gamma-Secretase Genes NCSTN, PSENEN, and PSEN1 Underlie Rare Forms of Hidradenitis Suppurativa (Acne Inversa). *The Journal of investigative dermatology* 132, 2459-2461.
23. Barth, J.H., Layton, A.M., and Cunliffe, W.J. (1996). Endocrine factors in pre- and postmenopausal women with hidradenitis suppurativa. *The British journal of dermatology* 134, 1057-1059.
24. Canoui-Poitrine, F., Revuz, J.E., Wolkenstein, P., Viallette, C., Gabison, G., Pouget, F., Poli, F., Faye, O., and Bastuji-Garin, S. (2009). Clinical characteristics of a series of 302 French patients with hidradenitis suppurativa, with an analysis of factors associated with disease severity. *Journal of the American Academy of Dermatology* 61, 51-57.

25. Lapins, J., Ye, W., Nyren, O., and Emtestam, L. (2001). Incidence of cancer among patients with hidradenitis suppurativa. *Archives of dermatology* 137, 730-734.
26. Sabat, R., Chanwangpong, A., Schneider-Burrus, S., Metternich, D., Kokolakis, G., Kurek, A., Philipp, S., Uribe, D., Wolk, K., and Sterry, W. (2012). Increased prevalence of metabolic syndrome in patients with acne inversa. *PloS one* 7, e31810.
27. Von Laffert, M., Stadie, V., Wohlrab, J., and Marsch, W.C. (2011). Hidradenitis suppurativa/acne inversa: bilocated epithelial hyperplasia with very different sequelae. *The British journal of dermatology* 164, 367-371.
28. Kamp, S., Fiehn, A.M., Stenderup, K., Rosada, C., Pakkenberg, B., Kemp, K., Dam, T.N., and Jemec, G.B. (2011). Hidradenitis suppurativa: a disease of the absent sebaceous gland? Sebaceous gland number and volume are significantly reduced in uninvolved hair follicles from patients with hidradenitis suppurativa. *The British journal of dermatology* 164, 1017-1022.
29. Revuz, J. (2009). Hidradenitis suppurativa. *Journal of the European Academy of Dermatology and Venereology* 23, 985-998.
30. Canoui-Poitaine, F., Le Thuaut, A., Revuz, J.E., Viallette, C., Gabison, G., Poli, F., Pouget, F., Wolkenstein, P., and Bastuji-Garin, S. (2013). Identification of three hidradenitis suppurativa phenotypes: latent class analysis of a cross-sectional study. *The Journal of investigative dermatology* 133, 1506-1511.
31. Wise, C.A., Gillum, J.D., Seidman, C.E., Lindor, N.M., Veile, R., Bashardes, S., and Lovett, M. (2002). Mutations in CD2BP1 disrupt binding to PTP PEST and are responsible for PAPA syndrome, an autoinflammatory disorder. *Human molecular genetics* 11, 961-969.
32. Marzano, A.V., Trevisan, V., Gattorno, M., Ceccherini, I., De Simone, C., and Crosti, C. (2013). Pyogenic Arthritis, Pyoderma Gangrenosum, Acne, and Hidradenitis Suppurativa (PAPASH): A New Autoinflammatory Syndrome Associated With a Novel Mutation of the PSTPIP1 Gene. *JAMA Dermatol* 149, 762-764.
33. Braun-Falco, M., Kovnerystyy, O., Lohse, P., and Ruzicka, T. (2012). Pyoderma gangrenosum, acne, and suppurative hidradenitis (PASH)--a new autoinflammatory syndrome distinct from PAPA syndrome. *Journal of the American Academy of Dermatology* 66, 409-415.
34. Kahn, M.F., and Chamot, A.M. (1992). SAPHO syndrome. *Rheumatic diseases clinics of North America* 18, 225-246.
35. Bergeron, J.R., and Stone, O.J. (1967). Interstitial keratitis associated with hidradenitis suppurativa. *Archives of dermatology* 95, 473-475.
36. Leybishkis, B., Fasseas, P., Ryan, K.F., and Roy, R. (2001). Hidradenitis suppurativa and acne conglobata associated with spondyloarthritis. *The American journal of the medical sciences* 321, 195-197.

37. Alzaga Fernandez, A.G., Demirci, H., Darnley-Fisch, D.A., and Steen, D.W. (2010). Interstitial keratitis secondary to severe hidradenitis suppurativa: a case report and literature review. *Cornea* 29, 1189-1191.
38. Blanco, R., Gonzalez-Vela, M.C., Gonzalez-Lopez, M.A., Fernandez-Llaca, H., Canal, J., and Gonzalez-Gay, M.A. (2012). Interstitial keratitis secondary to severe hidradenitis suppurativa responding to adalimumab. *Cornea* 31, 206; author reply 206.
39. Von Der Werth, J.M., Williams, H.C., and Raeburn, J.A. (2000). The clinical genetics of hidradenitis suppurativa revisited. *The British journal of dermatology* 142, 947-953.
40. Hurley, H.J. (1989). Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In *Dermatological Surgery*, Roenigk R.K., Roenigk H.H., eds (New York: Marcel Dekker), pp 729-739.
41. Sartorius, K., Lapins, J., Emtestam, L., and Jemec, G.B. (2003). Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. *The British journal of dermatology* 149, 211-213.
42. Finlay, A.Y., and Khan, G.K. (1994). Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clinical and experimental dermatology* 19, 210-216.
43. Mebazaa, A., Ben Hadid, R., Cheikh Rouhou, R., Trojjet, S., El Euch, D., Mokni, M., Zitouna, M., and Ben Osman, A. (2009). Hidradenitis suppurativa: a disease with male predominance in Tunisia. *Acta dermatovenerologica Alpina, Panonica, et Adriatica* 18, 165-172.
44. Wolkenstein, P., Loundou, A., Barrau, K., Auquier, P., and Revuz, J. (2007). Quality of life impairment in hidradenitis suppurativa: a study of 61 cases. *Journal of the American Academy of Dermatology* 56, 621-623.
45. Von der Werth, J.M., and Jemec, G.B. (2001). Morbidity in patients with hidradenitis suppurativa. *The British journal of dermatology* 144, 809-813.
46. Jemec, G.B., Heidenheim, M., and Nielsen, N.H. (1996). Hidradenitis suppurativa - characteristics and consequences. *Clinical and experimental dermatology* 21, 419-423.
47. Clemmensen, O.J. (1983). Topical treatment of hidradenitis suppurativa with clindamycin. *International journal of dermatology* 22, 325-328.
48. Jemec, G.B., and Wendelboe, P. (1998). Topical clindamycin versus systemic tetracycline in the treatment of hidradenitis suppurativa. *Journal of the American Academy of Dermatology* 39, 971-974.
49. Gener, G., Canoui-Poittrine, F., Revuz, J.E., Faye, O., Poli, F., Gabison, G., Pouget, F., Viallette, C., Wolkenstein, P., and Bastuji-Garin, S. (2009).

- Combination therapy with clindamycin and rifampicin for hidradenitis suppurativa: a series of 116 consecutive patients. *Dermatology* 219, 148-154.
50. Van der Zee, H.H., Boer, J., Prens, E.P., and Jemec, G.B. (2009). The effect of combined treatment with oral clindamycin and oral rifampicin in patients with hidradenitis suppurativa. *Dermatology* 219, 143-147.
  51. Boer, J., and Nazary, M. (2011). Long-term results of acitretin therapy for hidradenitis suppurativa. Is acne inversa also a misnomer? *The British journal of dermatology* 164, 170-175.
  52. Boer, J., and van Gemert, M.J. (1999). Long-term results of isotretinoin in the treatment of 68 patients with hidradenitis suppurativa. *Journal of the American Academy of Dermatology* 40, 73-76.
  53. Mortimer, P.S., Dawber, R.P., Gales, M.A., and Moore, R.A. (1986). A double-blind controlled cross-over trial of cyproterone acetate in females with hidradenitis suppurativa. *The British journal of dermatology* 115, 263-268.
  54. Sawers, R.S., Randall, V.A., and Ebling, F.J. (1986). Control of hidradenitis suppurativa in women using combined antiandrogen (cyproterone acetate) and oestrogen therapy. *The British journal of dermatology* 115, 269-274.
  55. Joseph, M.A., Jayaseelan, E., Ganapathi, B., and Stephen, J. (2005). Hidradenitis suppurativa treated with finasteride. *The Journal of dermatological treatment* 16, 75-78.
  56. Yazdanyar, S., Boer, J., Ingvarsson, G., Szepietowski, J.C., and Jemec, G.B. (2011). Dapsone therapy for hidradenitis suppurativa: a series of 24 patients. *Dermatology* 222, 342-346.
  57. Rose, R.F., Goodfield, M.J., and Clark, S.M. (2006). Treatment of recalcitrant hidradenitis suppurativa with oral ciclosporin. *Clinical and experimental dermatology* 31, 154-155.
  58. Van der Zee, H.H., and Prens, E.P. (2011). The anti-inflammatory drug colchicine lacks efficacy in hidradenitis suppurativa. *Dermatology* 223, 169-173.
  59. Jemec, G.B. (2002). Methotrexate is of limited value in the treatment of hidradenitis suppurativa. *Clinical and experimental dermatology* 27, 528-529.
  60. Blok, J.L., van Hattem, S., Jonkman, M.F., and Horvath, B. (2013). Systemic therapy with immunosuppressive agents and retinoids in hidradenitis suppurativa: a systematic review. *The British journal of dermatology* 168, 243-252.
  61. Frohlich, D., Baaske, D., and Glatzel, M. (2000). [Radiotherapy of hidradenitis suppurativa - still valid today?]. *Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft [et al]* 176, 286-289.

62. Trombetta, M., Werts, E.D., and Parda, D. (2010). The role of radiotherapy in the treatment of hidradenitis suppurativa: case report and review of the literature. *Dermatology online journal* 16, 16.
63. Tierney, E., Mahmoud, B.H., Hexsel, C., Ozog, D., and Hamzavi, I. (2009). Randomized control trial for the treatment of hidradenitis suppurativa with a neodymium-doped yttrium aluminium garnet laser. *Dermatologic surgery : official publication for American Society for Dermatologic Surgery [et al]* 35, 1188-1198.
64. Lapins, J., Marcusson, J.A., and Emtestam, L. (1994). Surgical treatment of chronic hidradenitis suppurativa: CO2 laser stripping-secondary intention technique. *The British journal of dermatology* 131, 551-556.
65. Schweiger, E.S., Riddle, C.C., and Aires, D.J. (2011). Treatment of hidradenitis suppurativa by photodynamic therapy with aminolevulinic acid: preliminary results. *Journal of drugs in dermatology : JDD* 10, 381-386.
66. Plewig, G., and Kligman, A. (1975). *Acne. Morphogenesis and Treatment.*(Berlin: Springer).
67. Scheinfeld, N.S. (2003). A case of dissecting cellulitis and a review of the literature. *Dermatology online journal* 9, 8.
68. Chicarilli, Z.N. (1987). Follicular occlusion triad: hidradenitis suppurativa, acne conglobata, and dissecting cellulitis of the scalp. *Annals of plastic surgery* 18, 230-237.
69. Conway, H., Stark, R.B., Climo, S., Weeter, J.C., and Garcia, F.A. (1952). The surgical treatment of chronic hidradenitis suppurativa. *Surgery, gynecology & obstetrics* 95, 455-464.
70. Bedlow, A.J., and Mortimer, P.S. (1996). Dowling-Degos disease associated with hidradenitis suppurativa. *Clinical and experimental dermatology* 21, 305-306.
71. Betz, R.C., Planko, L., Eigelshoven, S., Hanneken, S., Pasternack, S.M., Bussow, H., Van Den Bogaert, K., Wenzel, J., Braun-Falco, M., Rutten, A., et al. (2006). Loss-of-function mutations in the keratin 5 gene lead to Dowling-Degos disease. *American journal of human genetics* 78, 510-519.
72. Todd, P., Garioch, J., Rademaker, M., Susskind, W., Gemell, C., and Thomson, J. (1990). Pachyonychia congenita complicated by hidradenitis suppurativa: a family study. *The British journal of dermatology* 123, 663-666.
73. McLean, W.H., Rugg, E.L., Lunny, D.P., Morley, S.M., Lane, E.B., Swensson, O., Dopping-Hepenstal, P.J., Griffiths, W.A., Eady, R.A., Higgins, C., et al. (1995). Keratin 16 and keratin 17 mutations cause pachyonychia congenita. *Nature genetics* 9, 273-278.
74. Sharma, V.M., and Stein, S.L. (2007). A novel mutation in K6b in pachyonychia congenita type 2. *The Journal of investigative dermatology* 127, 2060-2062.



75. Montgomery, J.R., White, T.W., Martin, B.L., Turner, M.L., and Holland, S.M. (2004). A novel connexin 26 gene mutation associated with features of the keratitis-ichthyosis-deafness syndrome and the follicular occlusion triad. *Journal of the American Academy of Dermatology* 51, 377-382.
76. Spiller, R.F., and Knox, J.M. (1958). Fox-Fordyce disease with hidradenitis suppurativa. *The Journal of investigative dermatology* 31, 127-135.
77. Cooney, R., and Jewell, D. (2009). The genetic basis of inflammatory bowel disease. *Dig Dis* 27, 428-442.
78. Lees, C.W., and Satsangi, J. (2009). Genetics of inflammatory bowel disease: implications for disease pathogenesis and natural history. *Expert review of gastroenterology & hepatology* 3, 513-534.
79. Quaglietta, L., te Velde, A., Staiano, A., Troncone, R., and Hommes, D.W. (2007). Functional consequences of NOD2/CARD15 mutations in Crohn disease. *Journal of pediatric gastroenterology and nutrition* 44, 529-539.
80. Danby, F.W., Jemec, G.B., Marsch, W., and von Laffert, M. (2013). Preliminary findings suggest hidradenitis suppurativa may be due to defective follicular support. *The British journal of dermatology* 168, 1034-1039.
81. Van der Zee, H.H., de Ruiter, L., Boer, J., van den Broecke, D.G., den Hollander, J.C., Laman, J.D., and Prens, E.P. (2012). Alterations in leucocyte subsets and histomorphology in normal-appearing perilesional skin and early and chronic hidradenitis suppurativa lesions. *The British journal of dermatology* 166, 98-106.
82. Van der Zee, H.H., de Ruiter, L., van den Broecke, D.G., Dik, W.A., Laman, J.D., and Prens, E.P. (2011). Elevated levels of tumour necrosis factor (TNF)-alpha, interleukin (IL)-1beta and IL-10 in hidradenitis suppurativa skin: a rationale for targeting TNF-alpha and IL-1beta. *The British journal of dermatology* 164, 1292-1298.
83. Van Rappard, D.C., Leenarts, M.F., Meijerink-van 't Oost, L., and Mekkes, J.R. (2011). Comparing treatment outcome of infliximab and adalimumab in patients with severe hidradenitis suppurativa. *The Journal of dermatological treatment* 23, 284-289.
84. Ahmed, A.A., Nordlind, K., Schultzberg, M., and Liden, S. (1995). Immunohistochemical localization of IL-1 alpha-, IL-1 beta-, IL-6- and TNF-alpha-like immunoreactivities in human apocrine glands. *Archives of dermatological research* 287, 764-766.
85. Matusiak, L., Bieniek, A., and Szepietowski, J.C. (2009). Increased serum tumour necrosis factor-alpha in hidradenitis suppurativa patients: is there a basis for treatment with anti-tumour necrosis factor-alpha agents? *Acta dermato-venereologica* 89, 601-603.
86. Marzano, A.V., Cugno, M., Trevisan, V., Fanoni, D., Venegoni, L., Berti, E., and Crosti, C. (2010). Role of inflammatory cells, cytokines and matrix

metalloproteinases in neutrophil-mediated skin diseases. *Clinical and experimental immunology* 162, 100-107.

87. Hunger, R.E., Surovy, A.M., Hassan, A.S., Braathen, L.R., and Yawalkar, N. (2008). Toll-like receptor 2 is highly expressed in lesions of acne inversa and colocalizes with C-type lectin receptor. *The British journal of dermatology* 158, 691-697.
88. Dreno, B., Khammari, A., Brocard, A., Moyse, D., Blouin, E., Guillet, G., Leonard, F., and Knol, A.C. (2012). Hidradenitis suppurativa: the role of deficient cutaneous innate immunity. *Archives of dermatology* 148, 182-186.
89. Lapins, J., Asman, B., Gustafsson, A., Bergstrom, K., and Emtestam, L. (2001). Neutrophil-related host response in hidradenitis suppurativa: a pilot study in patients with inactive disease. *Acta dermato-venereologica* 81, 96-99.
90. Giamarellos-Bourboulis, E.J., Antonopoulou, A., Petropoulou, C., Mouktaroudi, M., Spyridaki, E., Baziaka, F., Pelekanou, A., Giamarellou, H., and Stavrianeas, N.G. (2007). Altered innate and adaptive immune responses in patients with hidradenitis suppurativa. *The British journal of dermatology* 156, 51-56.
91. Wolk, K., Warszawska, K., Hoeflich, C., Witte, E., Schneider-Burrus, S., Witte, K., Kunz, S., Buss, A., Roewert, H.J., Krause, M., et al. (2011). Deficiency of IL-22 contributes to a chronic inflammatory disease: pathogenetic mechanisms in acne inversa. *J Immunol* 186, 1228-1239.
92. Van der Zee, H.H., van der Woude, C.J., Florencia, E.F., and Prens, E.P. (2010). Hidradenitis suppurativa and inflammatory bowel disease: are they associated? Results of a pilot study. *The British journal of dermatology* 162, 195-197.
93. Hiemstra, I.H., Bouma, G., Geerts, D., Kraal, G., and den Haan, J.M. (2012). Nod2 improves barrier function of intestinal epithelial cells via enhancement of TLR responses. *Molecular immunology* 52, 264-272.
94. Rubino, S.J., Selvanantham, T., Girardin, S.E., and Philpott, D.J. (2012). Nod-like receptors in the control of intestinal inflammation. *Current opinion in immunology* 24, 398-404.
95. Glaser, R., Harder, J., Lange, H., Bartels, J., Christophers, E., and Schroder, J.M. (2005). Antimicrobial psoriasin (S100A7) protects human skin from *Escherichia coli* infection. *Nature immunology* 6, 57-64.
96. Braff, M.H., Hawkins, M.A., Di Nardo, A., Lopez-Garcia, B., Howell, M.D., Wong, C., Lin, K., Streib, J.E., Dorschner, R., Leung, D.Y., et al. (2005). Structure-function relationships among human cathelicidin peptides: dissociation of antimicrobial properties from host immunostimulatory activities. *J Immunol* 174, 4271-4278.
97. Schroder, J.M. (2011). Antimicrobial peptides in healthy skin and atopic dermatitis. *Allergology international : official journal of the Japanese Society of Allergology* 60, 17-24.

98. Chen, X., Niyonsaba, F., Ushio, H., Hara, M., Yokoi, H., Matsumoto, K., Saito, H., Nagaoka, I., Ikeda, S., Okumura, K., et al. (2007). Antimicrobial peptides human beta-defensin (hBD)-3 and hBD-4 activate mast cells and increase skin vascular permeability. *European journal of immunology* 37, 434-444.
99. Schlapbach, C., Yawalkar, N., and Hunger, R.E. (2009). Human beta-defensin-2 and psoriasin are overexpressed in lesions of acne inversa. *Journal of the American Academy of Dermatology* 61, 58-65.
100. Hofmann, S.C., Saborowski, V., Lange, S., Kern, W.V., Bruckner-Tuderman, L., and Rieg, S. (2011). Expression of innate defense antimicrobial peptides in hidradenitis suppurativa. *Journal of the American Academy of Dermatology*.
101. Bechara, F.G., Sand, M., Skrygan, M., Kreuter, A., Altmeyer, P., and Gambichler, T. (2012). Acne inversa: evaluating antimicrobial peptides and proteins. *Annals of dermatology* 24, 393-397.
102. Emelianov, V.U., Bechara, F.G., Glaser, R., Langan, E.A., Taungjaruwainai, W.M., Schroder, J.M., Meyer, K.C., and Paus, R. (2012). Immunohistological pointers to a possible role for excessive cathelicidin (LL-37) expression by apocrine sweat glands in the pathogenesis of hidradenitis suppurativa/acne inversa. *The British journal of dermatology* 166, 1023-1034.
103. Christoph, T., Muller-Rover, S., Audring, H., Tobin, D.J., Hermes, B., Cotsarelis, G., Ruckert, R., and Paus, R. (2000). The human hair follicle immune system: cellular composition and immune privilege. *The British journal of dermatology* 142, 862-873.
104. Meyer, K.C., Klatte, J.E., Dinh, H.V., Harries, M.J., Reithmayer, K., Meyer, W., Sinclair, R., and Paus, R. (2008). Evidence that the bulge region is a site of relative immune privilege in human hair follicles. *The British journal of dermatology* 159, 1077-1085.
105. Nagao, K., Ginhoux, F., Leitner, W.W., Motegi, S., Bennett, C.L., Clausen, B.E., Merad, M., and Udey, M.C. (2009). Murine epidermal Langerhans cells and langerin-expressing dermal dendritic cells are unrelated and exhibit distinct functions. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3312-3317.
106. Igyarto, B.Z., Haley, K., Ortner, D., Bobr, A., Gerami-Nejad, M., Edelson, B.T., Zurawski, S.M., Malissen, B., Zurawski, G., Berman, J., et al. (2011). Skin-resident murine dendritic cell subsets promote distinct and opposing antigen-specific T helper cell responses. *Immunity* 35, 260-272.
107. Nagao, K., Kobayashi, T., Moro, K., Ohyama, M., Adachi, T., Kitashima, D.Y., Ueha, S., Horiuchi, K., Tanizaki, H., Kabashima, K., et al. (2012). Stress-induced production of chemokines by hair follicles regulates the trafficking of dendritic cells in skin. *Nature immunology* 13, 744-752.
108. Heath, W.R., and Mueller, S.N. (2012). Hair follicles: gatekeepers to the epidermis. *Nature immunology* 13, 715-717.

109. Cornbleet, T. (1952). Pregnancy and apocrine gland diseases: hidradenitis, Fox-Fordyce disease. *AMA archives of dermatology and syphilology* 65, 12-19.
110. Mortimer, P.S., Dawber, R.P., Gales, M.A., and Moore, R.A. (1986). Mediation of hidradenitis suppurativa by androgens. *Br Med J (Clin Res Ed)* 292, 245-248.
111. Jemec, G.B. (1988). The symptomatology of hidradenitis suppurativa in women. *The British journal of dermatology* 119, 345-350.
112. Jain, S.K., Kannan, K., and Prouty, L. (2004). Progesterone, but not 17beta-estradiol, increases TNF-alpha secretion in U937 monocytes. *Cytokine* 26, 102-105.
113. Lai, J.J., Lai, K.P., Chuang, K.H., Chang, P., Yu, I.C., Lin, W.J., and Chang, C. (2009). Monocyte/macrophage androgen receptor suppresses cutaneous wound healing in mice by enhancing local TNF-alpha expression. *The Journal of clinical investigation* 119, 3739-3751.
114. Barth, J.H., and Kealey, T. (1991). Androgen metabolism by isolated human axillary apocrine glands in hidradenitis suppurativa. *The British journal of dermatology* 125, 304-308.
115. Alikhan, A., Lynch, P.J., and Eisen, D.B. (2009). Hidradenitis suppurativa: a comprehensive review. *Journal of the American Academy of Dermatology* 60, 539-561; quiz 562-533.
116. Jemec, G.B., Faber, M., Gutschik, E., and Wendelboe, P. (1996). The bacteriology of hidradenitis suppurativa. *Dermatology* 193, 203-206.
117. Sartorius, K., Killasli, H., Oprica, C., Sullivan, A., and Lapins, J. (2011). Bacteriology of hidradenitis suppurativa exacerbations. Deep tissue cultures obtained during carbon dioxide laser treatment. *The British journal of dermatology* 166, 879-883.
118. Lapins, J., Jarstrand, C., and Emtestam, L. (1999). Coagulase-negative staphylococci are the most common bacteria found in cultures from the deep portions of hidradenitis suppurativa lesions, as obtained by carbon dioxide laser surgery. *The British journal of dermatology* 140, 90-95.
119. Wortsman, X., Revuz, J., and Jemec, G.B. (2009). Lymph nodes in hidradenitis suppurativa. *Dermatology* 219, 22-24.
120. Ghanim, H., Aljada, A., Hofmeyer, D., Syed, T., Mohanty, P., and Dandona, P. (2004). Circulating mononuclear cells in the obese are in a proinflammatory state. *Circulation* 110, 1564-1571.
121. Huang, S., Rutkowski, J.M., Snodgrass, R.G., Ono-Moore, K.D., Schneider, D.A., Newman, J.W., Adams, S.H., and Hwang, D.H. (2012). Saturated fatty acids activate TLR-mediated proinflammatory signaling pathways. *Journal of lipid research* 53, 2002-2013.

122. Faergemann, J., Aly, R., Wilson, D.R., and Maibach, H.I. (1983). Skin occlusion: effect on *Pityrosporum orbiculare*, skin PCO<sub>2</sub>, pH, transepidermal water loss, and water content. *Archives of dermatological research* 275, 383-387.
123. Hana, A., Booken, D., Henrich, C., Gratchev, A., Maas-Szabowski, N., Goerdts, S., and Kurzen, H. (2007). Functional significance of non-neuronal acetylcholine in skin epithelia. *Life sciences* 80, 2214-2220.
124. Parks, R.W., and Parks, T.G. (1997). Pathogenesis, clinical features and management of hidradenitis suppurativa. *Annals of the Royal College of Surgeons of England* 79, 83-89.
125. Jeong, S.H., Park, J.H., Kim, J.N., Park, Y.H., Shin, S.Y., Lee, Y.H., Kye, Y.C., and Son, S.W. (2010). Up-regulation of TNF- $\alpha$  secretion by cigarette smoke is mediated by Egr-1 in HaCaT human keratinocytes. *Experimental dermatology* 19, e206-212.
126. Mosely, L.H., and Finseth, F. (1977). Cigarette smoking: impairment of digital blood flow and wound healing in the hand. *The Hand* 9, 97-101.
127. Leow, Y.H., and Maibach, H.I. (1998). Cigarette smoking, cutaneous vasculature, and tissue oxygen. *Clinics in dermatology* 16, 579-584.
128. Sorensen, L.T., Nielsen, H.B., Kharazmi, A., and Gottrup, F. (2004). Effect of smoking and abstention on oxidative burst and reactivity of neutrophils and monocytes. *Surgery* 136, 1047-1053.
129. Wong, L.S., Green, H.M., Feugate, J.E., Yadav, M., Nothnagel, E.A., and Martins-Green, M. (2004). Effects of "second-hand" smoke on structure and function of fibroblasts, cells that are critical for tissue repair and remodeling. *BMC cell biology* 5, 13.
130. Arredondo, J., Hall, L.L., Ndoye, A., Nguyen, V.T., Chernyavsky, A.I., Bercovich, D., Orr-Urtreger, A., Beaudet, A.L., and Grando, S.A. (2003). Central role of fibroblast  $\alpha 3$  nicotinic acetylcholine receptor in mediating cutaneous effects of nicotine. *Laboratory investigation; a journal of technical methods and pathology* 83, 207-225.
131. Yin, L., Morita, A., and Tsuji, T. (2000). Alterations of extracellular matrix induced by tobacco smoke extract. *Archives of dermatological research* 292, 188-194.
132. Fitzsimmons, J.S., and Guilbert, P.R. (1985). A family study of hidradenitis suppurativa. *Journal of medical genetics* 22, 367-373.
133. Lapins, J., Olerup, O., and Emtestam, L. (2001). No human leukocyte antigen-A, -B or -DR association in Swedish patients with hidradenitis suppurativa. *Acta dermato-venereologica* 81, 28-30.
134. Gao, M., Wang, P.G., Cui, Y., Yang, S., Zhang, Y.H., Lin, D., Zhang, K.Y., Liang, Y.H., Sun, L.D., Yan, K.L., et al. (2006). Inversa acne (hidradenitis

- suppurativa): a case report and identification of the locus at chromosome 1p21.1-1q25.3. *The Journal of investigative dermatology* 126, 1302-1306.
135. Wang, B., Yang, W., Wen, W., Sun, J., Su, B., Liu, B., Ma, D., Lv, D., Wen, Y., Qu, T., et al. (2010). Gamma-secretase gene mutations in familial acne inversa. *Science* 330, 1065.
  136. Pink, A.E., Simpson, M.A., Brice, G.W., Smith, C.H., Desai, N., Mortimer, P.S., Barker, J.N., and Trembath, R.C. (2011). PSENEN and NCSTN mutations in familial hidradenitis suppurativa (Acne Inversa). *The Journal of investigative dermatology* 131, 1568-1570.
  137. Liu, Y., Gao, M., Lv, Y.M., Yang, X., Ren, Y.Q., Jiang, T., Zhang, X., Guo, B.R., Li, M., Zhang, Q., et al. (2011). Confirmation by exome sequencing of the pathogenic role of NCSTN mutations in acne inversa (hidradenitis suppurativa). *The Journal of investigative dermatology* 131, 1570-1572.
  138. Li, C.R., Jiang, M.J., Shen, D.B., Xu, H.X., Wang, H.S., Yao, X., Zhang, Y., Zhou, W.Q., and Wang, B. (2011). Two novel mutations of the nicastrin gene in Chinese patients with acne inversa. *The British journal of dermatology* 165, 415-418.
  139. Miskinyte, S., Nassif, A., Merabtene, F., Ungeheuer, M.N., Join-Lambert, O., Jais, J.P., and Hovnanian, A. (2012). Nicastrin Mutations in French Families with Hidradenitis Suppurativa. *The Journal of investigative dermatology* 132, 1728-1730.
  140. Zhang, C., Wang, L., Chen, L., Ren, W., Mei, A., Chen, X., and Deng, Y. (2012). Two novel mutations of the NCSTN gene in Chinese familial acne inverse. *Journal of the European Academy of Dermatology and Venereology*, 27, 1571-1574.
  141. Nomura, Y., Nomura, T., Sakai, K., Sasaki, K., Ohguchi, Y., Mizuno, O., Hata, H., Aoyagi, S., Abe, R., Itaya, Y., et al. (2012). A novel splice site mutation in NCSTN underlies a Japanese family with hidradenitis suppurativa. *The British journal of dermatology* 168, 206-209.
  142. Jiao, T., Dong, H., Jin, L., Wang, S., and Wang, J. (2013). A novel nicastrin mutation in a large Chinese family with hidradenitis suppurativa. *The British journal of dermatology* 168, 1141-1143.
  143. Chavez-Gutierrez, L. (2013). Dissecting gamma-secretase function. *Journal of neurochemistry* 125, 1-3.
  144. Tolia, A., and De Strooper, B. (2009). Structure and function of gamma-secretase. *Seminars in cell & developmental biology* 20, 211-218.
  145. Kimberly, W.T., LaVoie, M.J., Ostaszewski, B.L., Ye, W., Wolfe, M.S., and Selkoe, D.J. (2003). Gamma-secretase is a membrane protein complex comprised of presenilin, nicastrin, Aph-1, and Pen-2. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6382-6387.

146. Annaert, W., and De Strooper, B. (2002). A cell biological perspective on Alzheimer's disease. *Annual review of cell and developmental biology* 18, 25-51.
147. Struhl, G., and Adachi, A. (2000). Requirements for presenilin-dependent cleavage of notch and other transmembrane proteins. *Molecular cell* 6, 625-636.
148. Hemming, M.L., Elias, J.E., Gygi, S.P., and Selkoe, D.J. (2008). Proteomic profiling of gamma-secretase substrates and mapping of substrate requirements. *PLoS biology* 6, e257.
149. Kopan, R., and Ilagan, M.X. (2004). Gamma-secretase: proteasome of the membrane? *Nature reviews Molecular cell biology* 5, 499-504.
150. Schenk, D. (2000). Alzheimer's disease. A partner for presenilin. *Nature* 407, 34-35.
151. Khandelwal, A., Chandu, D., Roe, C.M., Kopan, R., and Quatrano, R.S. (2007). Moonlighting activity of presenilin in plants is independent of gamma-secretase and evolutionarily conserved. *Proceedings of the National Academy of Sciences of the United States of America* 104, 13337-13342.
152. McMains, V.C., Myre, M., Kreppel, L., and Kimmel, A.R. (2010). Dictyostelium possesses highly diverged presenilin/gamma-secretase that regulates growth and cell-fate specification and can accurately process human APP: a system for functional studies of the presenilin/gamma-secretase complex. *Disease models & mechanisms* 3, 581-594.
153. Jutras, I., Laplante, A., Boulais, J., Brunet, S., Thinakaran, G., and Desjardins, M. (2005). Gamma-secretase is a functional component of phagosomes. *The Journal of biological chemistry* 280, 36310-36317.
154. Neely, K.M., Green, K.N., and LaFerla, F.M. (2011). Presenilin is necessary for efficient proteolysis through the autophagy-lysosome system in a gamma-secretase-independent manner. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31, 2781-2791.
155. Tamboli, I.Y., Prager, K., Thal, D.R., Thelen, K.M., Dewachter, I., Pietrzik, C.U., St George-Hyslop, P., Sisodia, S.S., De Strooper, B., Heneka, M.T., et al. (2008). Loss of gamma-secretase function impairs endocytosis of lipoprotein particles and membrane cholesterol homeostasis. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28, 12097-12106.
156. Vetrivel, K.S., Zhang, Y.W., Xu, H., and Thinakaran, G. (2006). Pathological and physiological functions of presenilins. *Molecular neurodegeneration* 1, 4.
157. Lichtenthaler, S.F., Haass, C., and Steiner, H. (2011). Regulated intramembrane proteolysis--lessons from amyloid precursor protein processing. *Journal of neurochemistry* 117, 779-796.
158. Bergmans, B.A., and De Strooper, B. (2010). gamma-secretases: from cell biology to therapeutic strategies. *Lancet neurology* 9, 215-226.

159. Sato, T., Diehl, T.S., Narayanan, S., Funamoto, S., Ihara, Y., De Strooper, B., Steiner, H., Haass, C., and Wolfe, M.S. (2007). Active gamma-secretase complexes contain only one of each component. *The Journal of biological chemistry* 282, 33985-33993.
160. Hebert, S.S., Serneels, L., Dejaegere, T., Horre, K., Dabrowski, M., Baert, V., Annaert, W., Hartmann, D., and De Strooper, B. (2004). Coordinated and widespread expression of gamma-secretase in vivo: evidence for size and molecular heterogeneity. *Neurobiology of disease* 17, 260-272.
161. De Strooper, B. (2003). Aph-1, Pen-2, and Nicastrin with Presenilin generate an active gamma-Secretase complex. *Neuron* 38, 9-12.
162. Mao, G., Cui, M.Z., Li, T., Jin, Y., and Xu, X. (2012). Pen-2 is dispensable for endoproteolysis of presenilin 1, and nicastrin-Aph subcomplex is important for both gamma-secretase assembly and substrate recruitment. *Journal of neurochemistry* 123, 837-844.
163. Francis, R., McGrath, G., Zhang, J., Ruddy, D.A., Sym, M., Apfeld, J., Nicoll, M., Maxwell, M., Hai, B., Ellis, M.C., et al. (2002). aph-1 and pen-2 are required for Notch pathway signaling, gamma-secretase cleavage of betaAPP, and presenilin protein accumulation. *Developmental cell* 3, 85-97.
164. Takasugi, N., Tomita, T., Hayashi, I., Tsuruoka, M., Niimura, M., Takahashi, Y., Thinakaran, G., and Iwatsubo, T. (2003). The role of presenilin cofactors in the gamma-secretase complex. *Nature* 422, 438-441.
165. Edbauer, D., Winkler, E., Haass, C., and Steiner, H. (2002). Presenilin and nicastrin regulate each other and determine amyloid beta-peptide production via complex formation. *Proceedings of the National Academy of Sciences of the United States of America* 99, 8666-8671.
166. De Strooper, B., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., Von Figura, K., and Van Leuven, F. (1998). Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* 391, 387-390.
167. Futai, E., Yagishita, S., and Ishiura, S. (2009). Nicastrin is dispensable for gamma-secretase protease activity in the presence of specific presenilin mutations. *The Journal of biological chemistry* 284, 13013-13022.
168. Ahn, K., Shelton, C.C., Tian, Y., Zhang, X., Gilchrist, M.L., Sisodia, S.S., and Li, Y.M. (2010). Activation and intrinsic gamma-secretase activity of presenilin 1. *Proceedings of the National Academy of Sciences of the United States of America* 107, 21435-21440.
169. Sherrington, R., Rogaev, E.I., Liang, Y., Rogaeva, E.A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K., et al. (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375, 754-760.



170. Spasic, D., and Annaert, W. (2008). Building gamma-secretase: the bits and pieces. *Journal of cell science* 121, 413-420.
171. Yu, G., Nishimura, M., Arawaka, S., Levitan, D., Zhang, L., Tandon, A., Song, Y.Q., Rogaeva, E., Chen, F., Kawarai, T., et al. (2000). Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and betaAPP processing. *Nature* 407, 48-54.
172. Goutte, C., Tsunozaki, M., Hale, V.A., and Priess, J.R. (2002). APH-1 is a multipass membrane protein essential for the Notch signaling pathway in *Caenorhabditis elegans* embryos. *Proceedings of the National Academy of Sciences of the United States of America* 99, 775-779.
173. Thinakaran, G., Borchelt, D.R., Lee, M.K., Slunt, H.H., Spitzer, L., Kim, G., Ratovitsky, T., Davenport, F., Nordstedt, C., Seeger, M., et al. (1996). Endoproteolysis of presenilin 1 and accumulation of processed derivatives in vivo. *Neuron* 17, 181-190.
174. Podlisny, M.B., Citron, M., Amarante, P., Sherrington, R., Xia, W., Zhang, J., Diehl, T., Levesque, G., Fraser, P., Haass, C., et al. (1997). Presenilin proteins undergo heterogeneous endoproteolysis between Thr291 and Ala299 and occur as stable N- and C-terminal fragments in normal and Alzheimer brain tissue. *Neurobiology of disease* 3, 325-337.
175. Hutton, M., and Hardy, J. (1997). The presenilins and Alzheimer's disease. *Human molecular genetics* 6, 1639-1646.
176. Yamasaki, A., Eimer, S., Okochi, M., Smialowska, A., Kaether, C., Baumeister, R., Haass, C., and Steiner, H. (2006). The GxGD motif of presenilin contributes to catalytic function and substrate identification of gamma-secretase. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26, 3821-3828.
177. Kretner, B., Fukumori, A., Kuhn, P.H., Perez-Revuelta, B.I., Lichtenthaler, S.F., Haass, C., and Steiner, H. (2013). Important functional role of residue x of the presenilin GxGD protease active site motif for APP substrate cleavage specificity and substrate selectivity of gamma-secretase. *Journal of neurochemistry* 125, 144-156.
178. Chavez-Gutierrez, L., Bammens, L., Benilova, I., Vandersteen, A., Benurwar, M., Borgers, M., Lismont, S., Zhou, L., Van Cleynenbreugel, S., Esselmann, H., et al. (2012). The mechanism of gamma-Secretase dysfunction in familial Alzheimer disease. *The EMBO journal* 31, 2261-2274.
179. Kaether, C., Capell, A., Edbauer, D., Winkler, E., Novak, B., Steiner, H., and Haass, C. (2004). The presenilin C-terminus is required for ER-retention, nicastrin-binding and gamma-secretase activity. *The EMBO journal* 23, 4738-4748.
180. Wang, J., Beher, D., Nyborg, A.C., Shearman, M.S., Golde, T.E., and Goate, A. (2006). C-terminal PAL motif of presenilin and presenilin homologues required for normal active site conformation. *Journal of neurochemistry* 96, 218-227.

181. Sato, C., Takagi, S., Tomita, T., and Iwatsubo, T. (2008). The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the gamma-secretase. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28, 6264-6271.
182. Li, X., Dang, S., Yan, C., Gong, X., Wang, J., and Shi, Y. (2013). Structure of a presenilin family intramembrane aspartate protease. *Nature* 493, 56-61.
183. Tolia, A., Horre, K., and De Strooper, B. (2008). Transmembrane domain 9 of presenilin determines the dynamic conformation of the catalytic site of gamma-secretase. *The Journal of biological chemistry* 283, 19793-19803.
184. Esler, W.P., Kimberly, W.T., Ostaszewski, B.L., Ye, W., Diehl, T.S., Selkoe, D.J., and Wolfe, M.S. (2002). Activity-dependent isolation of the presenilin-gamma -secretase complex reveals nicastrin and a gamma substrate. *Proceedings of the National Academy of Sciences of the United States of America* 99, 2720-2725.
185. Annaert, W.G., Esselens, C., Baert, V., Boeve, C., Snellings, G., Cupers, P., Craessaerts, K., and De Strooper, B. (2001). Interaction with telencephalin and the amyloid precursor protein predicts a ring structure for presenilins. *Neuron* 32, 579-589.
186. Herreman, A., Van Gassen, G., Bentahir, M., Nyabi, O., Craessaerts, K., Mueller, U., Annaert, W., and De Strooper, B. (2003). gamma-Secretase activity requires the presenilin-dependent trafficking of nicastrin through the Golgi apparatus but not its complex glycosylation. *Journal of cell science* 116, 1127-1136.
187. Lee, M.K., Slunt, H.H., Martin, L.J., Thinakaran, G., Kim, G., Gandy, S.E., Seeger, M., Koo, E., Price, D.L., and Sisodia, S.S. (1996). Expression of presenilin 1 and 2 (PS1 and PS2) in human and murine tissues. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 16, 7513-7525.
188. Herreman, A., Hartmann, D., Annaert, W., Saftig, P., Craessaerts, K., Serneels, L., Umans, L., Schrijvers, V., Checler, F., Vanderstichele, H., et al. (1999). Presenilin 2 deficiency causes a mild pulmonary phenotype and no changes in amyloid precursor protein processing but enhances the embryonic lethal phenotype of presenilin 1 deficiency. *Proceedings of the National Academy of Sciences of the United States of America* 96, 11872-11877.
189. Dries, D.R., and Yu, G. (2008). Assembly, maturation, and trafficking of the gamma-secretase complex in Alzheimer's disease. *Current Alzheimer research* 5, 132-146.
190. Martoglio, B., and Golde, T.E. (2003). Intramembrane-cleaving aspartic proteases and disease: presenilins, signal peptide peptidase and their homologs. *Human molecular genetics* 12 Spec No 2, R201-206.
191. Zhang, H., Sun, S., Herreman, A., De Strooper, B., and Bezprozvanny, I. (2010). Role of presenilins in neuronal calcium homeostasis. *The Journal of*

neuroscience : the official journal of the Society for Neuroscience 30, 8566-8580.

192. Shah, S., Lee, S.F., Tabuchi, K., Hao, Y.H., Yu, C., LaPlant, Q., Ball, H., Dann, C.E., 3rd, Sudhof, T., and Yu, G. (2005). Nicastrin functions as a gamma-secretase-substrate receptor. *Cell* 122, 435-447.
193. Capell, A., Kaether, C., Edbauer, D., Shirotani, K., Merkl, S., Steiner, H., and Haass, C. (2003). Nicastrin interacts with gamma-secretase complex components via the N-terminal part of its transmembrane domain. *The Journal of biological chemistry* 278, 52519-52523.
194. De Strooper, B. (2005). Nicastrin: gatekeeper of the gamma-secretase complex. *Cell* 122, 318-320.
195. Li, T., Ma, G., Cai, H., Price, D.L., and Wong, P.C. (2003). Nicastrin is required for assembly of presenilin/gamma-secretase complexes to mediate Notch signaling and for processing and trafficking of beta-amyloid precursor protein in mammals. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 23, 3272-3277.
196. Siman, R., and Velji, J. (2003). Localization of presenilin-nicastrin complexes and gamma-secretase activity to the trans-Golgi network. *Journal of neurochemistry* 84, 1143-1153.
197. Murphy, M.P., Das, P., Nyborg, A.C., Rochette, M.J., Dodson, M.W., Loosbrock, N.M., Souder, T.M., McLendon, C., Merit, S.L., Piper, S.C., et al. (2003). Overexpression of nicastrin increases Abeta production. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 17, 1138-1140.
198. Goo, J.S., Kim, Y.B., Shim, S.B., Jee, S.W., Lee, S.H., Kim, J.E., Hwang, I.S., Lee, Y.J., Kwak, M.H., Lim, C.J., et al. (2013). Nicastrin Overexpression in Transgenic Mice Induces Aberrant Behavior and APP Processing. *Molecular neurobiology* 48, 232-243.
199. Shirotani, K., Edbauer, D., Capell, A., Schmitz, J., Steiner, H., and Haass, C. (2003). Gamma-secretase activity is associated with a conformational change of nicastrin. *The Journal of biological chemistry* 278, 16474-16477.
200. Chavez-Gutierrez, L., Tolia, A., Maes, E., Li, T., Wong, P.C., and de Strooper, B. (2008). Glu(332) in the Nicastrin ectodomain is essential for gamma-secretase complex maturation but not for its activity. *The Journal of biological chemistry* 283, 20096-20105.
201. Chen, F., Yu, G., Arawaka, S., Nishimura, M., Kawarai, T., Yu, H., Tandon, A., Supala, A., Song, Y.Q., Rogaeva, E., et al. (2001). Nicastrin binds to membrane-tethered Notch. *Nature cell biology* 3, 751-754.
202. Hayashi, I., Takatori, S., Urano, Y., Miyake, Y., Takagi, J., Sakata-Yanagimoto, M., Iwanari, H., Osawa, S., Morohashi, Y., Li, T., et al. (2012). Neutralization of

the gamma-secretase activity by monoclonal antibody against extracellular domain of nicastrin. *Oncogene* 31, 787-798.

203. Zhang, X., Hoey, R.J., Lin, G., Koide, A., Leung, B., Ahn, K., Dolios, G., Paduch, M., Ikeuchi, T., Wang, R., et al. (2012). Identification of a tetratricopeptide repeat-like domain in the nicastrin subunit of gamma-secretase using synthetic antibodies. *Proceedings of the National Academy of Sciences of the United States of America* 109, 8534-8539.
204. Ilaya, N.T., Evin, G., Masters, C.L., and Culvenor, J.G. (2004). Nicastrin expression in mouse peripheral tissues is not co-ordinated with presenilin and is high in muscle. *Journal of neurochemistry* 91, 230-237.
205. Pardossi-Piquard, R., Dunys, J., Giaime, E., Guillot-Sestier, M.V., St George-Hyslop, P., Checler, F., and Alves da Costa, C. (2009). p53-dependent control of cell death by nicastrin: lack of requirement for presenilin-dependent gamma-secretase complex. *Journal of neurochemistry* 109, 225-237.
206. Kim, S.H., and Sisodia, S.S. (2005). Evidence that the "NF" motif in transmembrane domain 4 of presenilin 1 is critical for binding with PEN-2. *The Journal of biological chemistry* 280, 41953-41966.
207. Prokop, S., Haass, C., and Steiner, H. (2005). Length and overall sequence of the PEN-2 C-terminal domain determines its function in the stabilization of presenilin fragments. *Journal of neurochemistry* 94, 57-62.
208. Kim, S.H., and Sisodia, S.S. (2005). A sequence within the first transmembrane domain of PEN-2 is critical for PEN-2-mediated endoproteolysis of presenilin 1. *The Journal of biological chemistry* 280, 1992-2001.
209. Niimura, M., Isoo, N., Takasugi, N., Tsuruoka, M., Ui-Tei, K., Saigo, K., Morohashi, Y., Tomita, T., and Iwatsubo, T. (2005). Aph-1 contributes to the stabilization and trafficking of the gamma-secretase complex through mechanisms involving intermolecular and intramolecular interactions. *The Journal of biological chemistry* 280, 12967-12975.
210. Pardossi-Piquard, R., Yang, S.P., Kanemoto, S., Gu, Y., Chen, F., Bohm, C., Sevalle, J., Li, T., Wong, P.C., Checler, F., et al. (2009). APH1 polar transmembrane residues regulate the assembly and activity of presenilin complexes. *The Journal of biological chemistry* 284, 16298-16307.
211. Chen, A.C., Guo, L.Y., Ostaszewski, B.L., Selkoe, D.J., and LaVoie, M.J. (2010). Aph-1 associates directly with full-length and C-terminal fragments of gamma-secretase substrates. *The Journal of biological chemistry* 285, 11378-11391.
212. Serneels, L., Dejaegere, T., Craessaerts, K., Horre, K., Jorissen, E., Tousseyn, T., Hebert, S., Coolen, M., Martens, G., Zwijsen, A., et al. (2005). Differential contribution of the three Aph1 genes to gamma-secretase activity in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 102, 1719-1724.

213. Ma, G., Li, T., Price, D.L., and Wong, P.C. (2005). APH-1a is the principal mammalian APH-1 isoform present in gamma-secretase complexes during embryonic development. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 25, 192-198.
214. Tomita, T., and Iwatsubo, T. (2013). Structural biology of presenilins and signal peptide peptidases. *The Journal of biological chemistry* 288, 14673-14680.
215. Gu, Y., Chen, F., Sanjo, N., Kawarai, T., Hasegawa, H., Duthie, M., Li, W., Ruan, X., Luthra, A., Mount, H.T., et al. (2003). APH-1 interacts with mature and immature forms of presenilins and nicastrin and may play a role in maturation of presenilin.nicastrin complexes. *The Journal of biological chemistry* 278, 7374-7380.
216. Kaether, C., Haass, C., and Steiner, H. (2006). Assembly, trafficking and function of gamma-secretase. *Neuro-degenerative diseases* 3, 275-283.
217. Rajendran, L., Honsho, M., Zahn, T.R., Keller, P., Geiger, K.D., Verkade, P., and Simons, K. (2006). Alzheimer's disease beta-amyloid peptides are released in association with exosomes. *Proceedings of the National Academy of Sciences of the United States of America* 103, 11172-11177.
218. Kaether, C., Schmitt, S., Willem, M., and Haass, C. (2006). Amyloid precursor protein and Notch intracellular domains are generated after transport of their precursors to the cell surface. *Traffic* 7, 408-415.
219. Pasternak, S.H., Bagshaw, R.D., Guiral, M., Zhang, S., Ackerley, C.A., Pak, B.J., Callahan, J.W., and Mahuran, D.J. (2003). Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. *The Journal of biological chemistry* 278, 26687-26694.
220. Fassler, M., Zocher, M., Klare, S., de la Fuente, A.G., Scheuermann, J., Capell, A., Haass, C., Valkova, C., Veerappan, A., Schneider, D., et al. (2010). Masking of transmembrane-based retention signals controls ER export of gamma-secretase. *Traffic* 11, 250-258.
221. Kaether, C., Scheuermann, J., Fassler, M., Zilow, S., Shirotani, K., Valkova, C., Novak, B., Kacmar, S., Steiner, H., and Haass, C. (2007). Endoplasmic reticulum retention of the gamma-secretase complex component Pen2 by Rer1. *EMBO reports* 8, 743-748.
222. Spasic, D., Raemaekers, T., Dillen, K., Declerck, I., Baert, V., Serneels, L., Fullekrug, J., and Annaert, W. (2007). Rer1p competes with APH-1 for binding to nicastrin and regulates gamma-secretase complex assembly in the early secretory pathway. *The Journal of cell biology* 176, 629-640.
223. Lazarov, V.K., Fraering, P.C., Ye, W., Wolfe, M.S., Selkoe, D.J., and Li, H. (2006). Electron microscopic structure of purified, active gamma-secretase reveals an aqueous intramembrane chamber and two pores. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6889-6894.

224. Osenkowski, P., Li, H., Ye, W., Li, D., Aeschbach, L., Fraering, P.C., Wolfe, M.S., and Selkoe, D.J. (2009). Cryoelectron microscopy structure of purified gamma-secretase at 12 Å resolution. *Journal of molecular biology* 385, 642-652.
225. Duplan, E., Sevalle, J., Viotti, J., Goiran, T., Bauer, C., Renbaum, P., Levy-Lahad, E., Gautier, C.A., Corti, O., Leroudier, N., et al. (2013). Parkin differently regulates presenilin-1 and presenilin-2 functions by direct control of their promoter transcription. *Journal of molecular cell biology* 5, 132-142.
226. Thathiah, A., Spittaels, K., Hoffmann, M., Staes, M., Cohen, A., Horre, K., Vanbrabant, M., Coun, F., Baekelandt, V., Delacourte, A., et al. (2009). The orphan G protein-coupled receptor 3 modulates amyloid-beta peptide generation in neurons. *Science* 323, 946-951.
227. Liu, X., Zhao, X., Zeng, X., Bossers, K., Swaab, D.F., Zhao, J., and Pei, G. (2013). beta-arrestin1 regulates gamma-secretase complex assembly and modulates amyloid-beta pathology. *Cell research* 23, 351-365.
228. Thathiah, A., Horre, K., Snellinx, A., Vandeweyer, E., Huang, Y., Ciesielska, M., De Kloe, G., Munck, S., and De Strooper, B. (2013). beta-arrestin 2 regulates Abeta generation and gamma-secretase activity in Alzheimer's disease. *Nature medicine* 19, 43-49.
229. Michaki, V., Guix, F.X., Vennekens, K., Munck, S., Dingwall, C., Davis, J.B., Townsend, D.M., Tew, K.D., Feiguin, F., De Strooper, B., et al. (2012). Down-regulation of the ATP-binding cassette transporter 2 (Abca2) reduces amyloid-beta production by altering Nicastrin maturation and intracellular localization. *The Journal of biological chemistry* 287, 1100-1111.
230. Hussain, I., Fabregue, J., Anderes, L., Ousson, S., Borlat, F., Eligert, V., Berger, S., Dimitrov, M., Alattia, J.R., Fraering, P.C., et al. (2013). The role of gamma-secretase activating protein (GSAP) and imatinib in the regulation of gamma-secretase activity and amyloid-beta generation. *The Journal of biological chemistry* 288, 2521-2531.
231. Frykman, S., Teranishi, Y., Hur, J.Y., Sandebring, A., Yamamoto, N.G., Ancarcona, M., Nishimura, T., Winblad, B., Bogdanovic, N., Schedin-Weiss, S., et al. (2012). Identification of two novel synaptic gamma-secretase associated proteins that affect amyloid beta-peptide levels without altering Notch processing. *Neurochemistry international* 61, 108-118.
232. Mo, J.S., Yoon, J.H., Hong, J.A., Kim, M.Y., Ann, E.J., Ahn, J.S., Kim, S.M., Baek, H.J., Lang, F., Choi, E.J., et al. (2012). Phosphorylation of nicastrin by SGK1 leads to its degradation through lysosomal and proteasomal pathways. *PloS one* 7, e37111.
233. Kapoor, A., Wang, B.J., Hsu, W.M., Chang, M.Y., Liang, S.M., and Liao, Y.F. (2013). Retinoic Acid-Elicited RARalpha/RXRalpha Signaling Attenuates Abeta Production by Directly Inhibiting gamma-Secretase-Mediated Cleavage of Amyloid Precursor Protein. *ACS chemical neuroscience* 4, 1093-1100.

234. Stemmer, N., Strekalova, E., Djogo, N., Ploger, F., Loers, G., Lutz, D., Buck, F., Michalak, M., Schachner, M., and Kleene, R. (2013). Generation of Amyloid-beta Is Reduced by the Interaction of Calreticulin with Amyloid Precursor Protein, Presenilin and Nicastrin. *PloS one* 8, e61299.
235. Wakabayashi, T., Craessaerts, K., Bammens, L., Bentahir, M., Borgions, F., Herdewijn, P., Staes, A., Timmerman, E., Vandekerckhove, J., Rubinstein, E., et al. (2009). Analysis of the gamma-secretase interactome and validation of its association with tetraspanin-enriched microdomains. *Nature cell biology* 11, 1340-1346.
236. Chen, F., Hasegawa, H., Schmitt-Ulms, G., Kwarai, T., Bohm, C., Katayama, T., Gu, Y., Sanjo, N., Glista, M., Rogaeva, E., et al. (2006). TMP21 is a presenilin complex component that modulates gamma-secretase but not epsilon-secretase activity. *Nature* 440, 1208-1212.
237. Zhou, S., Zhou, H., Walian, P.J., and Jap, B.K. (2005). CD147 is a regulatory subunit of the gamma-secretase complex in Alzheimer's disease amyloid beta-peptide production. *Proceedings of the National Academy of Sciences of the United States of America* 102, 7499-7504.
238. Tanabe, C., Maeda, T., Zou, K., Liu, J., Liu, S., Nakajima, T., and Komano, H. (2012). The ubiquitin ligase synoviolin up-regulates amyloid beta production by targeting a negative regulator of gamma-secretase, Rer1, for degradation. *The Journal of biological chemistry* 287, 44203-44211.
239. Jeon, A.H., Bohm, C., Chen, F., Huo, H., Ruan, X., Ren, C.H., Ho, K., Qamar, S., Mathews, P.M., Fraser, P.E., et al. (2013). Interactome analyses of mature gamma-secretase complexes reveals distinct molecular environments of presenilin (PS) paralogs and preferential binding of signal peptide peptidase to PS2. *The Journal of biological chemistry* 288, 15352-15366.
240. Thinakaran, G., and Koo, E.H. (2008). Amyloid precursor protein trafficking, processing, and function. *The Journal of biological chemistry* 283, 29615-29619.
241. Vassar, R. (2004). BACE1: the beta-secretase enzyme in Alzheimer's disease. *Journal of molecular neuroscience* : MN 23, 105-114.
242. Allinson, T.M., Parkin, E.T., Turner, A.J., and Hooper, N.M. (2003). ADAMs family members as amyloid precursor protein alpha-secretases. *Journal of neuroscience research* 74, 342-352.
243. Citron, M., Westaway, D., Xia, W., Carlson, G., Diehl, T., Levesque, G., Johnson-Wood, K., Lee, M., Seubert, P., Davis, A., et al. (1997). Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nature medicine* 3, 67-72.
244. Ho, A., and Sudhof, T.C. (2004). Binding of F-spondin to amyloid-beta precursor protein: a candidate amyloid-beta precursor protein ligand that modulates amyloid-beta precursor protein cleavage. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2548-2553.

245. Fortini, M.E. (2009). Notch signaling: the core pathway and its posttranslational regulation. *Developmental cell* 16, 633-647.
246. Struhl, G., and Adachi, A. (1998). Nuclear access and action of notch in vivo. *Cell* 93, 649-660.
247. Okochi, M., Steiner, H., Fukumori, A., Tanii, H., Tomita, T., Tanaka, T., Iwatsubo, T., Kudo, T., Takeda, M., and Haass, C. (2002). Presenilins mediate a dual intramembranous gamma-secretase cleavage of Notch-1. *The EMBO journal* 21, 5408-5416.
248. Clark R.F., Hutton, M., Fuldner, S., Froelich, S., Karran, E., Talbot, C., Crook, R., Lendon, C. Prihar, G., He, C. et al. (1995). The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. Alzheimer's Disease Collaborative Group. *Nature genetics* 11, 219-222.
249. Song, W., Nadeau, P., Yuan, M., Yang, X., Shen, J., and Yankner, B.A. (1999). Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. *Proceedings of the National Academy of Sciences of the United States of America* 96, 6959-6963.
250. Jarrett, J.T., and Lansbury, P.T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73, 1055-1058.
251. McGowan, E., Pickford, F., Kim, J., Onstead, L., Eriksen, J., Yu, C., Skipper, L., Murphy, M.P., Beard, J., Das, P., et al. (2005). Abeta42 is essential for parenchymal and vascular amyloid deposition in mice. *Neuron* 47, 191-199.
252. Saito, T., Suemoto, T., Brouwers, N., Sleegers, K., Funamoto, S., Mihira, N., Matsuba, Y., Yamada, K., Nilsson, P., Takano, J., et al. (2011). Potent amyloidogenicity and pathogenicity of Abeta43. *Nature neuroscience* 14, 1023-1032.
253. Duff, K., Eckman, C., Zehr, C., Yu, X., Prada, C.M., Perez-tur, J., Hutton, M., Buee, L., Harigaya, Y., Yager, D., et al. (1996). Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature* 383, 710-713.
254. Kuperstein, I., Broersen, K., Benilova, I., Rozenski, J., Jonckheere, W., Debulpaep, M., Vandersteen, A., Segers-Nolten, I., Van Der Werf, K., Subramaniam, V., et al. (2010). Neurotoxicity of Alzheimer's disease Abeta peptides is induced by small changes in the Abeta42 to Abeta40 ratio. *The EMBO journal* 29, 3408-3420.
255. Tanzi, R.E., and Bertram, L. (2005). Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell* 120, 545-555.
256. Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nature genetics* 30, 97-101.



257. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841-842.
258. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
259. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* 38, e164.
260. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467, 1061-1073.
261. Plagnol, V., Curtis, J., Epstein, M., Mok, K.Y., Stebbings, E., Grigoriadou, S., Wood, N.W., Hambleton, S., Burns, S.O., Thrasher, A.J., et al. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 28, 2747-2754.
262. Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* 3, 1101-1108.
263. Hills, A., Ahn, J.W., Donaghue, C., Thomas, H., Mann, K., and Ogilvie, C.M. (2010). MLPA for confirmation of array CGH results and determination of inheritance. *Molecular cytogenetics* 3, 19.
264. Holmes, O., Paturi, S., Ye, W., Wolfe, M.S., and Selkoe, D.J. (2012). Effects of membrane lipids on the activity and processivity of purified gamma-secretase. *Biochemistry* 51, 3565-3575.
265. Li, Y.M., Lai, M.T., Xu, M., Huang, Q., DiMuzio-Mower, J., Sardana, M.K., Shi, X.P., Yin, K.C., Shafer, J.A., and Gardell, S.J. (2000). Presenilin 1 is linked with gamma-secretase activity in the detergent solubilized state. *Proceedings of the National Academy of Sciences of the United States of America* 97, 6138-6143.
266. Quintero-Monzon, O., Martin, M.M., Fernandez, M.A., Cappello, C.A., Krzysiak, A.J., Osenkowski, P., and Wolfe, M.S. (2011). Dissociation between the processivity and total activity of gamma-secretase: implications for the mechanism of Alzheimer's disease-causing presenilin mutations. *Biochemistry* 50, 9023-9035.
267. Carter, M.S., Doskow, J., Morris, P., Li, S., Nhim, R.P., Sandstedt, S., and Wilkinson, M.F. (1995). A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *The Journal of biological chemistry* 270, 28995-29003.
268. Schouten, J.P., McElgunn, C.J., Waaijer, R., Zwiijnenburg, D., Diepvens, F., and Pals, G. (2002). Relative quantification of 40 nucleic acid sequences by

- multiplex ligation-dependent probe amplification. *Nucleic acids research* 30, e57.
269. Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* 102, 765-775.
  270. Atkinson, G.C., Baldauf, S.L., and Hauryliuk, V. (2008). Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components. *BMC evolutionary biology* 8, 290.
  271. Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440, 561-564.
  272. Filipovic, A., Gronau, J.H., Green, A.R., Wang, J., Vallath, S., Shao, D., Rasul, S., Ellis, I.O., Yague, E., Sturge, J., et al. (2011). Biological and clinical implications of nicastrin expression in invasive breast cancer. *Breast cancer research and treatment* 125, 43-53.
  273. Zhao, G., Liu, Z., Ilagan, M.X., and Kopan, R. (2010). Gamma-secretase composed of PS1/Pen2/Aph1a can cleave notch and amyloid precursor protein in the absence of nicastrin. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, 1648-1656.
  274. Nam, S.H., Seo, S.J., Goo, J.S., Kim, J.E., Choi, S.I., Lee, H.R., Hwang, I.S., Jee, S.W., Lee, S.H., Bae, C.J., et al. (2011). Pen-2 overexpression induces Abeta-42 production, memory defect, motor activity enhancement and feeding behavior dysfunction in NSE/Pen-2 transgenic mice. *International journal of molecular medicine* 28, 961-971.
  275. Haass, C., Schlossmacher, M.G., Hung, A.Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B.L., Lieberburg, I., Koo, E.H., Schenk, D., Teplow, D.B., et al. (1992). Amyloid beta-peptide is produced by cultured cells during normal metabolism. *Nature* 359, 322-325.
  276. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols* 4, 1073-1081.
  277. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. *Nature methods* 7, 248-249.
  278. Mucaki, E.J., Shirley, B.C., and Rogan, P.K. (2013). Prediction of mutant mRNA splice isoforms by information theory-based exon definition. *Human mutation* 34, 557-565.
  279. Ingram, J.R., Wood, M., John, B., Butler, R., and Anstey, A.V. (2013). Absence of pathogenic gamma-secretase mutations in a South Wales cohort of familial and sporadic hidradenitis suppurativa (acne inversa). *The British journal of dermatology* 168, 874-876.

280. Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461, 272-276.
281. Onoufriadis, A., Simpson, M.A., Pink, A.E., Di Meglio, P., Smith, C.H., Pullabhatla, V., Knight, J., Spain, S.L., Nestle, F.O., Burden, A.D., et al. (2011). Mutations in IL36RN/IL1F5 are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. *American journal of human genetics* 89, 432-437.
282. Bilguvar, K., Ozturk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoglu, D., Tuysuz, B., Caglayan, A.O., Gokben, S., et al. (2010). Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature* 467, 207-210.
283. Simpson, M.A., Irving, M.D., Asilmaz, E., Gray, M.J., Dafou, D., Elmslie, F.V., Mansour, S., Holder, S.E., Brain, C.E., Burton, B.K., et al. (2011). Mutations in NOTCH2 cause Hajdu-Cheney syndrome, a disorder of severe and progressive bone loss. *Nature genetics* 43, 303-305.
284. Ostergaard, P., Simpson, M.A., Connell, F.C., Steward, C.G., Brice, G., Woollard, W.J., Dafou, D., Kilo, T., Smithson, S., Lunt, P., et al. (2011). Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nature genetics* 43, 929-931.
285. Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237-241.
286. Ostergaard, P., Simpson, M.A., Brice, G., Mansour, S., Connell, F.C., Onoufriadis, A., Child, A.H., Hwang, J., Kalidas, K., Mortimer, P.S., et al. (2011). Rapid identification of mutations in GJC2 in primary lymphoedema using whole exome sequencing combined with linkage analysis with delineation of the phenotype. *Journal of medical genetics* 48, 251-255.
287. Sauna, Z.E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nature reviews Genetics* 12, 683-691.
288. Li, M., Cheng, R., Liang, J., Yan, H., Zhang, H., Yang, L., Li, C., Jiao, Q., Lu, Z., He, J., et al. (2013). Mutations in POFUT1, Encoding Protein O-fucosyltransferase 1, Cause Generalized Dowling-Degos Disease. *American journal of human genetics* 92, 895-903.
289. Loo, W.J., Rytina, E., and Todd, P.M. (2004). Hidradenitis suppurativa, Dowling-Degos and multiple epidermal cysts: a new follicular occlusion triad. *Clinical and experimental dermatology* 29, 622-624.

290. Pink, A.E., Simpson, M.A., Desai, N., Trembath, R.C., and Barker, J.N. (2013). gamma-Secretase mutations in hidradenitis suppurativa: new insights into disease pathogenesis. *The Journal of investigative dermatology* 133, 601-607.
291. Li, T., Wen, H., Brayton, C., Das, P., Smithson, L.A., Fauq, A., Fan, X., Crain, B.J., Price, D.L., Golde, T.E., et al. (2007). Epidermal growth factor receptor and notch pathways participate in the tumor suppressor function of gamma-secretase. *The Journal of biological chemistry* 282, 32264-32273.
292. Ingram, J.R., and Piguet, V. (2013). Phenotypic heterogeneity in hidradenitis suppurativa (acne inversa): classification is an essential step toward personalized therapy. *The Journal of investigative dermatology* 133, 1453-1456.
293. Panza, F., Frisardi, V., Imbimbo, B.P., Capurso, C., Logroscino, G., Sancarlo, D., Seripa, D., Vendemiale, G., Pilotto, A., and Solfrizzi, V. (2010). REVIEW: gamma-Secretase inhibitors for the treatment of Alzheimer's disease: The current state. *CNS neuroscience & therapeutics* 16, 272-284.
294. Kelleher, R.J., 3rd, and Shen, J. (2010). Genetics. Gamma-secretase and human disease. *Science* 330, 1055-1056.
295. Pan, Y., Lin, M.H., Tian, X., Cheng, H.T., Gridley, T., Shen, J., and Kopan, R. (2004). gamma-secretase functions through Notch signaling to maintain skin appendages but is not required for their patterning or initial morphogenesis. *Developmental cell* 7, 731-743.
296. Xia, X., Qian, S., Soriano, S., Wu, Y., Fletcher, A.M., Wang, X.J., Koo, E.H., Wu, X., and Zheng, H. (2001). Loss of presenilin 1 is associated with enhanced beta-catenin signaling and skin tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10863-10868.
297. Oswald, F., Tauber, B., Dobner, T., Bourteele, S., Kostezka, U., Adler, G., Liptay, S., and Schmid, R.M. (2001). p300 acts as a transcriptional coactivator for mammalian Notch-1. *Molecular and cellular biology* 21, 7761-7774.
298. Estrach, S., Legg, J., and Watt, F.M. (2007). Syntenin mediates Delta1-induced cohesiveness of epidermal stem cells in culture. *Journal of cell science* 120, 2944-2952.
299. Lowell, S., Jones, P., Le Roux, I., Dunne, J., and Watt, F.M. (2000). Stimulation of human epidermal differentiation by delta-notch signalling at the boundaries of stem-cell clusters. *Current biology : CB* 10, 491-500.
300. Watt, F.M., Estrach, S., and Ambler, C.A. (2008). Epidermal Notch signalling: differentiation, cancer and adhesion. *Current opinion in cell biology* 20, 171-179.
301. Lefort, K., Mandinova, A., Ostano, P., Kolev, V., Calpini, V., Kolfschoten, I., Devgan, V., Lieb, J., Raffoul, W., Hohl, D., et al. (2007). Notch1 is a p53 target gene involved in human keratinocyte tumor suppression through negative regulation of ROCK1/2 and MRCKalpha kinases. *Genes & development* 21, 562-577.

302. Zhang, Q., Wang, C., Liu, Z., Liu, X., Han, C., Cao, X., and Li, N. (2012). Notch signal suppresses Toll-like receptor-triggered inflammatory responses in macrophages by inhibiting extracellular signal-regulated kinase 1/2-mediated nuclear factor kappaB activation. *The Journal of biological chemistry* 287, 6208-6217.
303. Alam, M.S., Maekawa, Y., Kitamura, A., Tanigaki, K., Yoshimoto, T., Kishihara, K., and Yasutomo, K. (2010). Notch signaling drives IL-22 secretion in CD4+ T cells by stimulating the aryl hydrocarbon receptor. *Proceedings of the National Academy of Sciences of the United States of America* 107, 5943-5948.
304. Haraguchi, K., Suzuki, T., Koyama, N., Kumano, K., Nakahara, F., Matsumoto, A., Yokoyama, Y., Sakata-Yanagimoto, M., Masuda, S., Takahashi, T., et al. (2009). Notch activation induces the generation of functional NK cells from human cord blood CD34-positive cells devoid of IL-15. *J Immunol* 182, 6168-6178.
305. Tilley, A.E., Harvey, B.G., Heguy, A., Hackett, N.R., Wang, R., O'Connor, T.P., and Crystal, R.G. (2009). Down-regulation of the notch pathway in human airway epithelium in association with smoking and chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 179, 457-466.
306. Melnik, B.C., and Plewig, G. (2013). Impaired Notch-MKP-1 signalling in hidradenitis suppurativa: an approach to pathogenesis by evidence from translational biology. *Experimental dermatology* 22, 172-177.
307. Giliberto, L., Borghi, R., Piccini, A., Mangerini, R., Sorbi, S., Cirmena, G., Garuti, A., Ghetti, B., Tagliavini, F., Mughal, M.R., et al. (2009). Mutant presenilin 1 increases the expression and activity of BACE1. *The Journal of biological chemistry* 284, 9027-9038.
308. Park, M.H., Choi, D.Y., Jin, H.W., Yoo, H.S., Han, J.Y., Oh, K.W., Han, S.B., Hwang, D.Y., and Hong, J.T. (2012). Mutant Presenilin 2 Increases beta-Secretase Activity Through Reactive Oxygen Species-Dependent Activation of Extracellular Signal-Regulated Kinase. *Journal of neuropathology and experimental neurology* 71, 130-139.
309. Tu, H., Nelson, O., Bezprozvanny, A., Wang, Z., Lee, S.F., Hao, Y.H., Serneels, L., De Strooper, B., Yu, G., and Bezprozvanny, I. (2006). Presenilins form ER Ca<sup>2+</sup> leak channels, a function disrupted by familial Alzheimer's disease-linked mutations. *Cell* 126, 981-993.
310. Bezprozvanny, I., and Mattson, M.P. (2008). Neuronal calcium mishandling and the pathogenesis of Alzheimer's disease. *Trends in neurosciences* 31, 454-463.
311. Chau, D.M., Crump, C.J., Villa, J.C., Scheinberg, D.A., and Li, Y.M. (2012). Familial Alzheimer disease presenilin-1 mutations alter the active site conformation of gamma-secretase. *The Journal of biological chemistry* 287, 17288-17296.

312. Franberg, J., Svensson, A.I., Winblad, B., Karlstrom, H., and Frykman, S. (2011). Minor contribution of presenilin 2 for gamma-secretase activity in mouse embryonic fibroblasts and adult mouse brain. *Biochemical and biophysical research communications* 404, 564-568.

## Appendix 1 – All known gamma-secretase substrates

### All reported gamma-secretase substrates

Substrate	Function
Alcadeins	Cadherin-related membrane protein
APLP1-2	Synaptogenesis/ neurite outgrowth
ApoER2	Signalling receptor
APP	Unknown role, Alzheimers disease
Bri2	Unknown role
CD43	Cell adhesion
CD44	Cell adhesion
CSF-1	Protein-tyrosine kinase
CXCL16, CX3CL1	Chemokine ligands
DCC	Netrin-1 receptor
Delta	Notch receptor ligand
DNER	Notch receptor ligand
DSG2	Desmosomes
Dystroglycan	Dystrophin-glycoprotein complex
E-Cadherin	Cell adhesion
EphB2	Receptor tyrosine kinase
Ephrin-B1	Eph receptor ligand
Ephrin-B2	Receptor tyrosine kinase
Erb4	Receptor tyrosine kinase
Gamma-protocadherins	Cell adhesion
GHR	Growth hormone receptor
GluR subunit 3	Glutamate receptor
HLA-A2	MHC Class I protein
IFN $\alpha$ R2	IFN- $\alpha$ receptor
Insulin receptor	Receptor tyrosine kinase
IGFI-R	Receptor tyrosine kinase
Ire1 $\alpha$ / Ire1 $\beta$	ER TM protein
IL-1R II	Cytokine receptor
IL-1R I	Cytokine receptor
Jagged	Notch receptor ligand
LAR	Receptor tyrosine phosphatase
LDLR	Lipoprotein receptor
LRP	Signalling receptor
Megalin	Signalling receptor
N-Cadherin	Cell adhesion
NCAM-L1	Cell adhesion
Nectin-1 $\alpha$	Adherens junction protein
Notch 1-4	Signalling receptor
NPR-C	Natriuretic peptide receptor
NRADD	Death-receptor-like protein
NRG-1	Growth factor
P75	Neurotrophin receptor
PLXDC2	Nervous system protein
Polyductin	Cell receptor
RPTP	Cell adhesion receptor

SorLA	Sorting receptor
SorCS1-3	Sorting receptor
Sortilin	Sorting receptor
Syndecan1-3	Heparan sulphate proteoglycan co-receptor
Tie1	Receptor tyrosine kinase
Tyrosinase/ tyrosinase related proteins 1-2	Pigment synthesis
VGSC $\beta$ 1-4 subunits	Sodium channel
Vasorin	TGF- $\beta$ inhibitor
VE-Cadherin	Cell adhesion
VLDL receptor	Signalling receptor
VEGFR-1	Growth factor receptor



## **Appendix 2: Study protocol and documentation**

**PROTOCOL TITLE:**  
**Molecular Genetic Analysis of Hidradenitis Suppurativa**

Chief Investigator (or Principal)

Name: Professor Jonathan Barker  
Title: Consultant Dermatologist  
Address: Floor 9, Guy's Tower, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 020 718 86412  
Fax: 020 718 86113  
Email: jonathan.barker@kcl.ac.uk

Name and address of Co-Investigator(s), Statistician, Laboratories etc  
Main site (Guy's and St. Thomas' Hospitals, King's College London):

Name: Dr Andrew Pink  
Title: MRC Clinical Research Training Fellow  
Address: Floor 9, Guy's Tower, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 020 718 86412  
Fax: 020 718 86113  
Email: andrew.pink@kcl.ac.uk

Name: Dr. Nemesha Desai  
Title: Consultant Dermatologist  
Address: St. John's Institute of Dermatology, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 02071886412  
Fax: 02071886113  
Email: nemesha.desai@gstt.nhs.uk

Name: Dr Catherine Smith  
Title: Consultant Dermatologist  
Address: Floor 9, Guy's Tower, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 020 718 86412  
Fax: 020 718 86113  
Email: [andrew.pink@kcl.ac.uk](mailto:andrew.pink@kcl.ac.uk)

Name: Naomi Hare  
Title: Clinical Research Nurse  
Address: Skin Therapy Research Unit, St. Thomas' Hospital, London. SE1 7EH  
Telephone: 020 71886277  
Fax: 020 7188 6113  
Email: [Naomi.hare@gstt.nhs.uk](mailto:Naomi.hare@gstt.nhs.uk)

Name: Richard Trembath  
Title: Professor, Head of Division of Genetics and Molecular Medicine  
Address: Floor 8, Guy's Tower, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 020 71886412  
Fax: 020 7188 6113  
Email: [Richard.trembath@kcl.ac.uk](mailto:Richard.trembath@kcl.ac.uk)

Name: Michael Simpson  
Title: Post-doctoral fellow  
Address: Floor 8, Guy's Tower, Guy's Hospital, Great Maze Pond, London, SE1 9RT  
Telephone: 02071886412  
Fax: 02071886113

Email: Michael.simpson@kcl.ac.uk

Second site (St. George's Hospital):

Name: Professor Peter Mortimer

Address: St. George's, Blackshaw Road, Tooting London SW17 0QT

Telephone: 0208 725 1785

Fax: 0208 725 3297

Email: [mortimer@sgul.ac.uk](mailto:mortimer@sgul.ac.uk)

## 1. BACKGROUND AND RATIONALE

Hidradenitis suppurativa (HS), otherwise known as acne inversa, is a common cause of chronic inflammatory skin disease with a prevalence as high as 1 per 100 in European populations. It is characterised by painful cutaneous nodules, cysts and abscesses in flexural sites resulting in sinus formation, fibrosis, and scarring with the potential for malignant transformation (squamous cell carcinoma) [1]. It commonly results in significant physical, psychological and social morbidity.

The pathogenesis of this disease is poorly understood, such that studies to date have been unable to distinguish whether HS represents a primary disorder of the hair follicle, the apocrine gland or an intrinsic defect of the host response. The efficacy of anti-tumour necrosis factor- $\alpha$  therapy indicates that the condition is maintained by an intense inflammatory response [2]. Importantly, environmental factors including obesity and in particular smoking have been strongly implicated in disease development (up to 90% of sufferers smoke [3]).

There is substantial evidence for an inherited component in the aetiology of HS. Fitzsimmons *et al* [4] demonstrated that 34% of first degree relatives of probands were affected and reassessment of the same families 15 years later supported an autosomal dominant inheritance pattern. Until recently, three putative genetic loci had been proposed including a very large (80cM) region at chromosome 1p21.1-1q25.3 [5], a 0.53cM region at chromosome 6q25 and a 16.8cM region spanning the centromere of chromosome 19 (McLean, unpublished data), but no causative genes had been identified. Mutations have recently been reported in the gamma-secretase genes (PSENEN, PSEN1 and NCSTN) [5], which we have replicated and reported in two of seven families that we had recruited with Hidradenitis Suppurativa [6]. How mutations in those genes result in disease are unclear however it has been proposed that it maybe via the gamma-secretase-NOTCH signaling pathway [5]. It is interesting to note that of the seven families that we had recruited with hidradenitis suppurativa, only two had mutations in these genes, indicating likely genetic heterogeneity [6]. The relevance of mutations in the gamma-secretase genes in sporadic cases of hidradenitis suppurativa is as yet unknown.

In our own tertiary referral clinic at St John's Institute of Dermatology, London, we have observed that up to 40% of patients with HS describe a family history of the condition, whereas 60% appear to be sporadic cases. Via the recruitment of familial and sporadic cases of HS we intend to identify the other genes underlying this condition and determine the relevance of all known mutations in sporadic cases. We intend to use Sangar sequencing, genome-wide linkage analysis, exome sequencing and bio-informatics to identify new disease genes. The second aim of our study is to initiate functional studies in order to understand how mutations in the gamma-secretase genes, and any further genes identified, result in the HS phenotype. This work will require tissue from affected individuals, such as hair plucks and skin, in order to set up cell cultures to perform in-vitro cell work..

## **2. STUDY OBJECTIVES, DESIGN AND STATISTICS:**

### **2.1 Study Objectives**

Primary objectives:

1. To identify the genes underlying HS
2. To determine the relevance of those genes in sporadic cases
3. To understand how mutations in those genes result in the disease phenotype.

### **2.2 Study Design:**

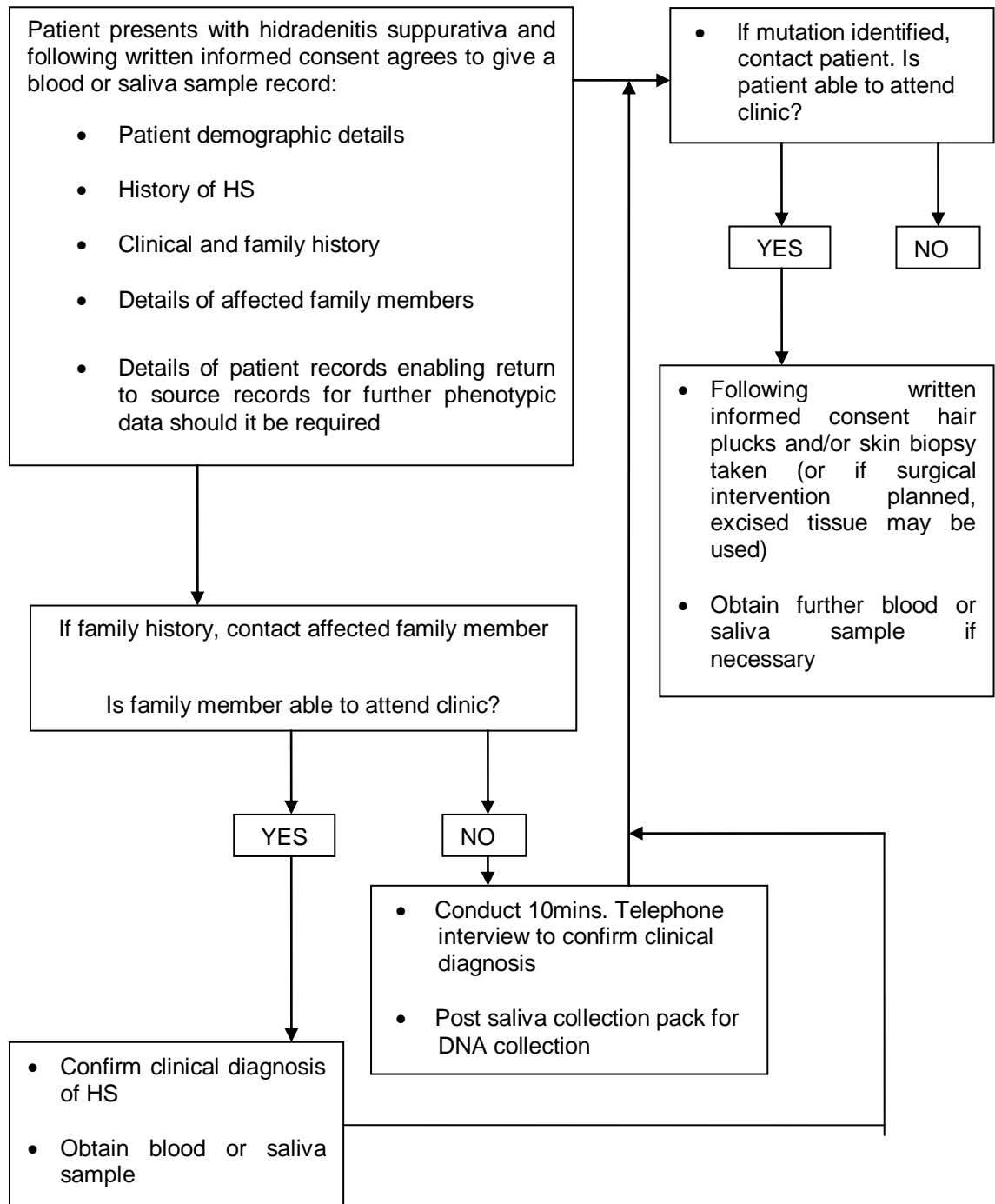
There are five main components of the study design:

1. Clinical questioning and examination to identify patients with HS (familial and sporadic cases) and collection of a blood or saliva sample for DNA extraction.
2. Clinical questioning +/- examination of relatives of patients with HS (only those with a family history) and collection of a blood or saliva sample for DNA extraction.
3. Genetic analysis of extracted DNA to identify candidate genes within families and to determine the relevance of those genes in sporadic cases.
4. In those patients in which mutations are identified, hair plucks and/or skin biopsies (or tissue from planned surgical interventions) will be performed to gain tissue for functional analysis.
5. Healthy volunteers will be recruited to provide blood, hair plucks and skin to act as normal controls.

### **2.3 Study Duration:**

Data and sample collection will take place over a maximum period of 5 years.

### **2.4 Study Flowchart:**



## 2.5 Study Statistics:

### 2.5.1 Sample size

A maximum of 50 families, 300 sporadic cases and 10 healthy control subjects will be recruited. These figures are based on the estimated number of recruitable patients seen in the primary centre (Guy's and St Thomas' Hospitals) in a 3 year period and on pilot data generated from sequencing the first 24 sporadic cases for mutations in the gamma-secretase genes. This sequencing revealed that 3 of 24 (12.5%) sporadic cases had mutations in the gamma-secretase genes (*NCSTN*, *PSEN1* and *PSENEN*).

The involvement of 300 sporadic cases will enable us to achieve a more accurate estimation of population frequency and would provide us with an estimated 37 patients with mutations for functional analysis. By recruiting 50 families we will have enough large pedigrees to successfully perform genome-wide linkage analysis and perform exome sequencing. Statistical analysis of genetic data will be performed using non-parametric or parametric statistical methods as appropriate in the Department of Genetics and Molecular Medicine, King's College London.

### **3. SELECTION AND WITHDRAWAL OF SUBJECTS:**

Target population:

Our target population will consist of patients attending our tertiary centre Hidradenitis Suppurativa clinic at St. John's Institute of Dermatology, Guy's and St. Thomas' Hospitals, London.

Inclusion Criteria:

1. Patient given written informed consent.
2. Patient 16 years or older
3. Patient with an established diagnosis of hidradenitis suppurativa (based on typical history and clinical examination)

Exclusion criteria

1. Patients who are unable to give written informed consent.
2. Patients aged under 16 years
3. Patients who have received a blood transfusion in the last 4 weeks
4. Patients in whom a diagnosis of hidradenitis suppurativa is not established.

Withdrawal of Subjects

Patients can withdraw from the study at any time. This will not affect his/her current or future treatment. The principal investigator/co-investigators can decide to withdraw a patient from the study at any time if they feel their participation is inappropriate.

### **4. INVESTIGATIONAL PLAN**

#### **4.1 Clinical Studies:**

There are four main components of the study design:

1. Clinical questioning and examination to identify patients with HS (familial and sporadic cases) and collection of a blood or saliva sample for DNA extraction.
2. Clinical questioning +/- examination of relatives of patients with HS (only those with a family history) and collection of a blood or saliva sample for DNA extraction.
3. Genetic analysis of extracted DNA to identify candidate genes within families and to determine the relevance of those genes in sporadic cases.
4. In those patients in which mutations are identified, hair plucks and/or skin biopsies will be performed to gain tissue for functional analysis. Alternatively, if a surgical procedure is planned, excised tissue may be used.
5. Healthy controls will be recruited so as to supply genetic material and tissue to act as a normal control.

## 1. Patients.

i. Patients attending dermatology clinic (referred by their GP, other dermatologists or hospital practitioners) for clinical assessment and treatment of their disease will be invited to participate in the study. Patients who agree to participate in the study will undertake detailed clinical genetic questioning as described below:

Date of birth  
Gender  
Recruitment source  
Weight  
Height  
Waist circumference  
Blood pressure  
Smoking status (current, past and pack-year history)  
Ethnicity  
History of 5 or more boils in flexural sites  
Sites affected with HS  
Age when first affected  
Relation of symptoms to menstrual cycle  
Past medical history  
Family history (+ draw pedigree if relevant)

These include questions to establish affected and unaffected members of their family. A family tree will be constructed, focussing on relatives with HS. Patients will also be asked to complete a Dermatology Life Quality Index score.

ii. Patients will be asked to provide a small sample of blood (17ml EDTA vacutainer) . This will be performed by a trained phlebotomist, a dermatologist or a dermatology nurse (who has been trained in phlebotomy during the course of their career). Patients who are needle phobic or unable to give blood will be asked to provide a 2ml sample of saliva using collection pots. Blood and saliva samples will be taken as described in Appendix 1.

iii. Blood and/ or saliva will be transported to St John's Institute of Dermatology (King's College London), Guys Hospital, London by courier or by post where it will be stored in the laboratory of the division of genetics and molecular medicine (storage of biological material will be in accordance with section 4.3; inter-site transport of samples will be in accordance with Appendix 1).

iv. Genetic studies on blood and or saliva samples will be carried out at St John's Institute of Dermatology (King's College), Guys Hospital, London.

v. Patients with affected family members will be asked to deliver study invitations and patient information leaflets to unaffected and affected family members or to provide contact details for family members.

## 2. Family members

i. Affected and unaffected family members will be invited to attend the department of dermatology for a short clinical interview to establish patient demographics and medical history (questions as above).

ii. Family members will be invited to have a brief clinical examination of the skin to confirm if they have a diagnosis of HS or a related disease.

iii. Family members who are unwilling/ unable to attend the department of dermatology will be invited to participate in a 10 minute telephone interview about skin symptoms to try and accurately establish if they are suffering with HS or related diseases.

iv.a. Family members able to attend the department will be asked to provide a small sample of blood (17ml EDTA vacutainer). This will be performed by a trained phlebotomist, a dermatologist or a dermatology nurse (who has been trained in phlebotomy during the course of their career).

iv.b. Family members unable to attend the department or those who are unable to donate blood (e.g. needle phobia) will be asked to donate a sample of saliva.

v. Blood or saliva will be transported to St John's Institute of Dermatology (King's College), Guys Hospital, London by courier or post where it will be stored in the laboratory of the Division of Genetics and Molecular Medicine.

vi. Blood or saliva will be stored at St John's Institute of Dermatology (King's College London), Guys Hospital, London, in the laboratory of the division of genetics and molecular medicine.

vii. Genetic studies on DNA extracted from the blood or saliva sample will be carried out at St John's Institute of Dermatology (King's College), Guys Hospital, London.

The inconvenience of hospital attendance for family members and minimal risk and discomfort of having a blood test and/ or providing a sample of saliva for patients and their families will be outweighed by the anticipated benefit for the community of HS patients. Although the research will not directly benefit participants during the course of the study, it may lead to the development of novel therapies in the future. The results of the research will be presented to peers at scientific meetings and in peer-reviewed journals

### 3. Affected individuals in whom mutations are identified.

i. Affected individuals in whom mutations are identified will be asked to re-attend the department for 6 hair plucks and/ or a skin biopsy. Alternatively, if surgical intervention is planned as a treatment for their disease then excised tissue may be used.

ii. 6 hairs will be plucked from the scalp as described in appendix 2. Two 6mm skin biopsies will be obtained from a lesional site and a non-lesional site in affected individuals. The skin biopsies will be performed using standard NHS protocols for biopsy sampling.

iii. Hair plucks and skin biopsies will be stored in accordance with section 4.3

### 4. Healthy control subjects

i. Healthy volunteers will be asked to attend the department and undertake questioning as detailed below:

Gender

Age

Ethnicity

Any history of hidradenitis suppurativa

Any history of related conditions such as acne, pilonidal abscesses, dissecting cellulitis



ii. If they report no history of those conditions they will be asked to provide a small blood sample (17ml, as detailed for patients above), 6 hairs will be plucked from the scalp as described in appendix 2 and two 6mm skin biopsies will be obtained from a flexural site (eg. Axilla) and a non-flexural site (eg. Back), performed using standard NHS protocols for biopsy sampling.

iii. Hair plucks and skin biopsies will be stored in accordance with section 4.3

#### 4.2 Sample DNA - Insufficient DNA concentration

In the event of there being too low a concentration of DNA in the sample the patient would be contacted by letter and asked to provide a further sample in the form of saliva (if the patient is not returning for a Dermatology clinic appointment in the near future (within 4 weeks)). The patient would initially be contacted by letter and asked to complete and return a reply slip. If the patient agrees they will be sent a saliva collection kit, asked to donate a sample and told to return it to the Skin Therapy Research Unit for further analysis.

#### 4.3 Laboratory studies:

DNA from affected and unaffected family members will be extracted from blood samples using standard methodologies. In participants unable to provide blood, DNA will be extracted from saliva samples.

Affected individuals will be screened for mutations in the gamma-secretase genes via standard Sanger sequencing methods. As well as determining which familial forms of the disease are due to mutations in these genes, this work will help to determine the genetic epidemiology of the disease by identifying how many sporadic cases are due to mutations in these genes.

The next part of the project will explore the hypothesis that there are more, as yet unidentified, genes of large effect size underlying familial forms of HS. This genetic heterogeneity is supported by our recent work demonstrating that only 2 of 7 pedigrees with HS had mutations in the gamma-secretase genes. We propose to use genome-wide linkage analysis in large families with no mutations in the gamma-secretase to identify further candidate genes underlying HS. Genetic loci giving strong statistical support for linked genes will then be submitted to positional cloning and candidate genes will be sequenced. One or more affected individuals from some of these, and smaller, families may be exome sequenced [7] (a technique that scans the whole exome of patients, capable of identifying novel mutations). By combining the results of these investigations along with bioinformatic analysis we aim to identify further genes underlying this debilitating condition. Validity of positive findings will be sought in further families and sporadic cases and also in patients with severe acne vulgaris. The latter are being collected on a large scale at St John's Institute of Dermatology (>2000 patients).

The final component of the project will explore the hypothesis that the pathogenic mechanisms of HS can be elucidated through the exploration of the biological role of the causative gene and its protein product. This work will involve the use of tissue harvested via skin biopsies and/ or hair plucks [8]. Experimentally, we will establish whether the restriction of the clinical phenotype to flexural sites is a consequence of a tissue specific expression pattern by generating gene and protein expression profiles across a range of tissues. These experiments will be extended with *in situ* hybridisation and immunohistological investigation to ascertain the cellular expression patterns in the hair follicle and the apocrine sweat glands. We will also undertake a series of experiments to delineate the cellular role of the protein product. We aim to define cellular phenotypes associated with over expression and gene silencing of the HS

gene. We will undertake these studies through the generation of gene expression and siRNA constructs and transfection into primary keratinocytes and HACAT cell lines. We will also explore the hypothesis that the mutation(s) identified disrupt the cellular phenotype associated with over expression by undertaking site-directed mutagenesis to incorporate the mutation into the expression construct. Model based systems are essential tools for establishing the biological mechanisms of a genetic disease. We aim to initiate the generation of a cellular model of HS through induced pluripotent stem cell techniques.

The potential benefits of this research are enormous with the possibility of diagnostic, prognostic and therapeutic advance led by the identification of underlying genes and a better understanding of their function.

#### Storage of Biological Specimens

Blood, saliva, skin and hair plucks will be stored in a secure facility at St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, King's College London, 9<sup>th</sup> Floor, Guy's Tower, Guy's Hospital, London, SE1 9RT.

Whole blood for DNA will be stored at -20°C until extraction. Saliva will be stored at ambient temperature until DNA extraction. Skin samples and hair plucks will be processed shortly after harvesting.

### **5. IMPLICATION OF RESULTS:**

Isolation of candidate genes for HS and a better understanding of their function will contribute towards an improved understanding of the pathogenesis and molecular mechanisms that underlie HS and related disorders.

### **6. SAFETY:**

#### Procedures for recording and Adverse Events

Although the occurrence of an adverse event is unlikely, the Principal Investigator is ultimately responsible for any adverse events that occur. Co-investigators, dermatologists and dermatology nurses will be available to assess and appropriately respond to such event e.g. the provision of first aid in the event of fainting during venepuncture.

### **7. PARTICIPANT RE-IMBURSEMENT**

Volunteers will be re-imbursed £150 to cover the inconvenience and the time they spend participating in the study. This is in respect of patient time and inconvenience regarding donation of blood samples, hair plucks, two skin biopsies and return to clinic for biopsy stitch removal 10 days later. Their initial visit will last approx 60 minutes and the follow-up appointment will last 10 minutes.

In addition reasonable travel costs will be reimbursed.

## **8. ADMINISTRATIVE ASPECTS**

### **8.1 Good Clinical practice**

The planning, conduct and reporting of this study will be in the spirit of the International conference on Harmonisation in Good Clinical Practice (ICH-GCP) 1996.

### **8.2 Declaration of Helsinki and Ethical Review**

The study will be performed in accordance with the principles stated in the Declaration of Helsinki, 2008.

Any protocol amendments will be submitted to the London Stanmore REC for approval (Research Ethics Committee Office, level 7, maternity block, Northwick Park Hospital, Watford Road, Harrow, HA1 3UJ). These will comply with ICH-GCP guidelines over the reporting of adverse events (AE's), serious adverse events (SAE's) & suspected, unexpected serious adverse reactions (SUSAR's) as well as providing the REC with progress reports, and a copy of the Final Study Report

### **8.3 Subject information and consent**

The investigator will ensure that the subject is given full and adequate verbal and written information about the nature, purpose, possible risk and benefit of the study. Subjects must also be notified that they are free to discontinue their participation in the study at any time. The subject will be given appropriate time to consider their participation in the study and have the opportunity to ask questions. Patient and family information leaflets are attached. The informed consent forms are also attached. Written informed consent will be obtained from all subjects before enrolment into the study. The subject

### **8.4 Subject protection**

Subjects will be assigned an anonymised study number to ensure subject confidentiality throughout the duration of the study. Only the subject number will be referenced in data summaries and study publications and presentations. The principal investigators will be responsible for keeping a Subject Identification Log of all subjects enrolled into the study and their corresponding study number. This information will be kept on a secure NHS server in a password protected file and will only be available to study personnel.

### **8.5 Direct access to source data and documents**

Key investigators of the project will be the only personnel with access to patient and study data and all data will be stored in accordance with the Data Protection Act, 1998. The chief investigator (Professor J Barker) will have overall control of and act as the custodian for all data for the full duration of time.

### **8.6 Peer review**

The study plan has been reviewed externally and approved to ensure the study is scientifically sound as part of the process of achieving funding through the Medical Research Council.

## **9. OTHER STUDY ISSUES**

### **9.1 Study timetable**

Estimated start date: August 2011

Estimated end date: August 2016

### **9.2 Changes to the present protocol**

The study must be conducted as defined by the present protocol. All changes must be documented by signed protocol amendments or a revised protocol, which will be submitted to the appropriate REC for approval. The investigator is responsible for notifying and obtaining approval from the Ethics Committee and the Regulatory Authorities for any changes to the protocol before implementation. National requirements will be followed.

### **9.3 Retention of study records**

Copies of protocols, results, correspondence, informed consents and other documents relevant to the study will be kept on file by the Principal Investigator and retained for at least 15 years after completion of discontinuation of the research study.

### **9.4 Publications**

The results of the study will be reported and disseminated in peer reviewed scientific journal and/ or conference presentations.

### **9.5 Study termination**

The study may be terminated at any time for reasons of safety and tolerability as determined by the principle investigator.

## **REFERENCES**

- 1) Malaguarnera M, et al. *Lancet* 1996; 348: 1449
- 2) Mekkes JR, et al. *Br J Dermatol* 2008; 158: 370-4
- 3) Konig A, et al. *Dermatology* 1999; 198: 261-4.
- 4) Fitzsimmons JS, et al. *J Med Genet* 1985; 22: 367-73.
- 5) Wang B, et al. *Science* 2010; 330: 1065
- 6) Pink et al. *JID* 2011 in press
- 7) Simpson MA et al. *Nat genet; epub March 2011*
- 8) Aasen T et al. *Nature protocols* 2010; 5: 371-382.

## Protocol Appendix 1

Sample collection and transport:

The patient can give either a saliva sample or a blood sample for collection of their DNA sample. Please note that a blood sample is preferable.

Collection and transport of Saliva Samples:

The saliva sample (2ml.) should be collected as per the manufacturer's instructions which can be found in the packaging of the pot. The patient must not have had anything to eat or drink for 30mins prior to giving the sample.

The pot must be labeled with the participant's site-specific study number using a cryopen and placed in a clear plastic specimen bag. A sample request form should accompany the saliva sample with details of the participant study number, date of birth, gender, ethnicity, date of sample collection, site name and name of dermatologist.

Saliva samples are stable at room temperature for up to 12 months and as such samples may be posted in batches of 6-10. Prior to postage please ensure the pot is sealed securely. Place the pot, specimen bag and form inside a Freepost padded envelope and post to the Skin Therapy Research Unit\* using normal post.

Collection and transport of Blood Sample:

Blood should be collected in accordance with site-specific local guidelines. Where appropriate the patient should be given a completed sample request form and collection tubes and be seen by the phlebotomist or if not available then by the research study nurse if appropriately trained.

Two pink 6ml. EDTA vacutainer tubes should be used for blood collection. Once collected blood tubes should be inverted 3x to ensure thorough mixing of blood and EDTA. These should be labeled using a cryo-pen with the participants surname and the date of sample collection and placed in a clear plastic specimen bag with a sample request form as detailed above.

Depending upon local site arrangements blood may be dispatched soon after collection or stored at -20°C

If facilities are available blood should be stored at either -20°C or at -80°C (depending upon equipment available to the Department) prior to collection by courier and transported on dry-ice to the Skin Therapy Research Unit. Collections can be arranged once there are a sufficient number of blood tubes to optimize the collection process. If freezer storage space is not available blood tubes should be placed in a pre-paid Royal Mail blood transport box and posted to the Skin Therapy Research Unit\* using normal post as soon as possible. Blood samples may be retained for up to 3 days at 4°C if available prior to postage if necessary to avoid posting on a Friday, or to maximize the number of tubes that are dispatched at any one time.

\*Please dispatch to the following address until further notice:

Skin Therapy Research Unit  
St. John's Institute of Dermatology  
St. Thomas's Hospital  
Westminster Bridge Road  
London SE1 7RT

## Protocol Appendix 2

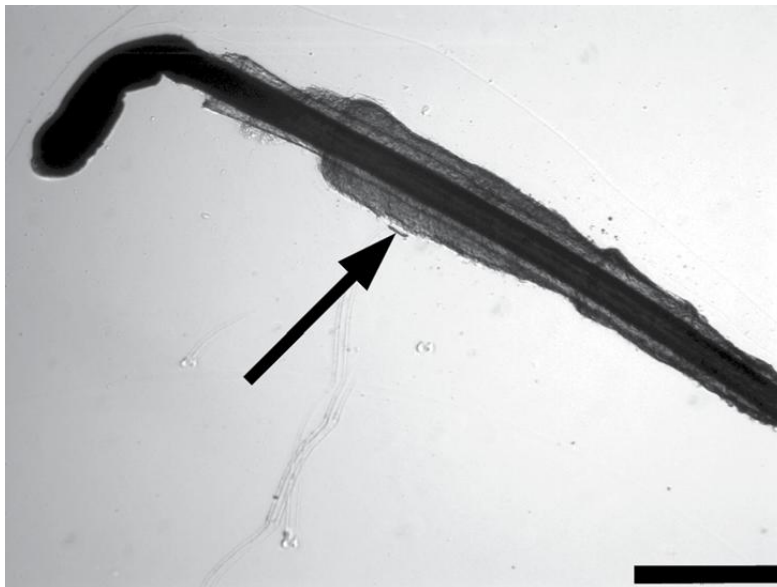
Protocol for the derivation of keratinocytes from plucked hair • TIMING 2-3 weeks

1| Prepare a non-coated 100-mm dish containing HBSS and Gentamycin 50 $\mu$ g/ml

2| Use tweezers to gently pull the hair out and place in HBSS medium. It is possible to use hair from any part of the body, although the occipital part of the head is particularly suitable and accessible, giving many hairs in anagen growth phase with a large amount of ORS cells on the hair shaft

**CRITICAL STEP** Although a noninvasive procedure, the donor should agree and sign consent forms as with normal biopsies.

**CRITICAL STEP** Hair should be in anagen growth phase with a clearly visible ORS (as shown below) and should be immediately transferred to some medium to avoid the drying out of the cells on the hair.



3| While submerged in the medium, cut off the external part of the hair leaving the bulb and ORS (one may also remove the bulb of the hair).

4| At this stage (according to the Belmonte lab), two optional procedures for growing keratinocytes from the plucked hair are described, direct outgrowth and enzymatic digestion.

5| Coat the required number of 6-well plates with Matrigel by adding sufficient Matrigel to cover the plate and incubate overnight at 4 °C.

6| Place the hair obtained from Step 2 in the coated culture plates.

**CRITICAL STEP** During the first 24 h, it is absolute essential to ensure that the hair sticks to the culture dish and does not float.

*Important.* In the Belmonte protocol they say to use MEF-coditioned hES media at this stage

7| Media change daily with MEF conditioned DMEM advanced (+10%serum, +bFGF, w/o antibiotics) and outgrowth is usually visible after 3-4 days.

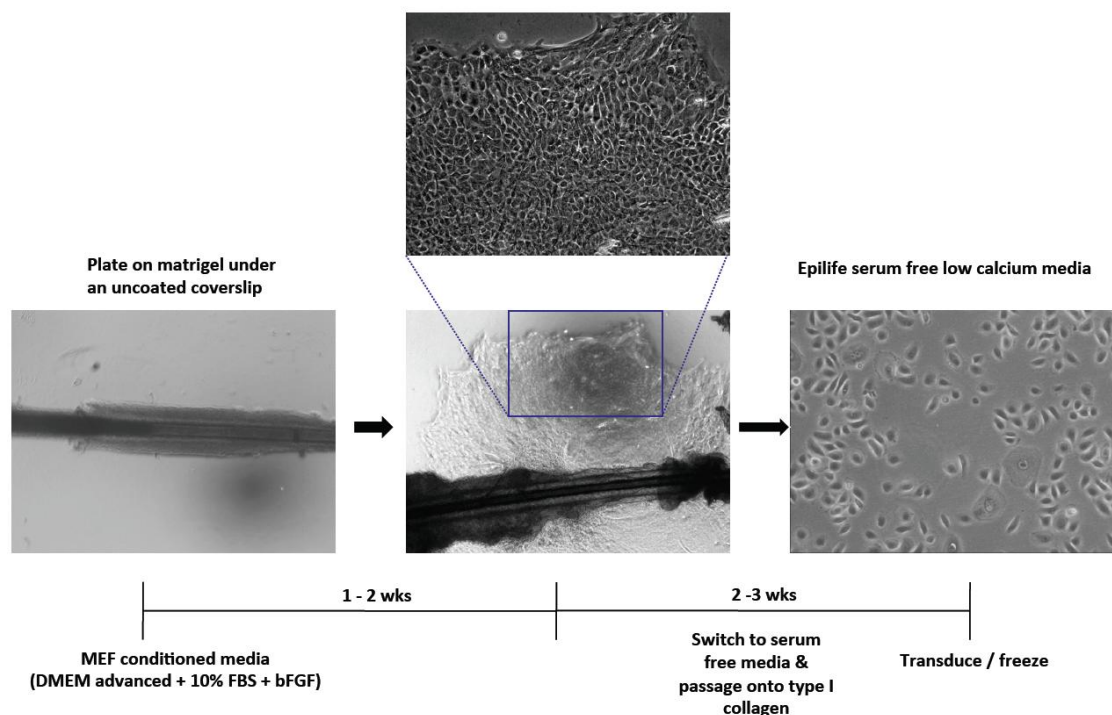
8| After around 7 days switch to Epilife based media (no adaptation necessary). The keratinocytes will then proliferate more rapidly and migrate.

9| After a further 7 days in Epilife media they are then usually ready to passage. The cells are very strongly adherent at this stage, especially around the hair shaft. It is therefore necessary to use 2.5% trypsin (rather than the usual 0.25%) to lift them off. I would also recommend not having serum in the trypsin neutralising solution as I find even a brief exposure to serum results in differentiation once the keratinocytes are adapted to low calcium Epilife.

10| From this stage onwards the cells are ready to either freeze down in small aliquots or transduce.

#### Method overview

#### **Deriving keratinocytes from scalp hairs.**



## **Genetics Studies of Hidradenitis Suppurativa**

Dear

We would like to invite you to take part in our research study of the skin disease hidradenitis suppurativa. Please find enclosed a patient information leaflet detailing what the study involves and why we would like your help.

With best wishes

Dr Andrew Pink  
Dermatology Registrar

Professor Jonathan Barker  
Professor of Dermatological Medicine



## **Genetics Studies of Hidradenitis Suppurativa**

Dear .....,

I am writing to you because your relative, ..... has a confirmed diagnosis of hidradenitis suppurativa and has attended the specialist Hidradenitis Suppurativa clinic at Guy's Hospital, London. He/She has given your name as a relative who may or may not be diagnosed with hidradenitis suppurativa and I am writing to ask if you would consider participating in a research study looking at the genetics of this condition. I have included a patient information leaflet for your consideration. We are interested in all family members, including anyone with 'normal' skin.

We are carrying out a research project at St. Thomas' Hospital to look at the genes that may cause hidradenitis suppurativa and to identify a target for novel therapies in the future. If at all possible we would like to see you in clinic to confirm if you have this diagnosis. This would be done by a trained dermatologist who would check your skin. We would then ask you if you would kindly donate a sample of blood or saliva.

If you are unable to attend the clinic could we conduct a short interview over the telephone, which would not last longer than 10mins. We can send you a saliva collection pack for a sample of your DNA so you would not have to attend for a blood sample.

If you are interested in participating I would ask that you return the reply slip so we may contact you at a time that is convenient for you.

You are under no obligation to take part in this study, and if you decline to take part, this will not affect any medical care you might be receiving at the moment or need in the future. If you do not wish to take part I would ask you to return the reply slip ticking the box that you are not interested, that way you will not be approached further.

Please see the attached supplementary information leaflet for some more information about this research.

Best wishes and many thanks

Yours sincerely

Dr Andrew Pink

Yours sincerely

Prof. Jonathan Barker

**Genetics Studies of Hidradenitis Suppurativa**

**REPLY SLIP**

Patient name: .....

Date of birth: .....

I **do not** wish to take part in this research, please do not contact me further – ☐

I **am** interested in taking part in the study ☐

Please indicate if you are able to attend at clinic.

Able/not able\* to attend - \*delete as applicable

Please indicate a time that would be suitable to you for us to conduct a short telephone interview if you are un-able to attend at clinic.

Date:

Time:

Current Address:.....

.....  
.....  
.....  
.....  
.....

Tel Number: .....

Email: .....

Many Thanks.

## **Information Leaflet: Patient**

### **Genetics Studies of Hidradenitis Suppurativa**

#### **Our invitation**

You are being invited to take part in a research study. Before you decide it is important to understand why the research is being done and what it will involve. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

#### **What is the purpose of the study?**

Hidradenitis Suppurativa (HS) is a chronic skin disease in which patients suffer with recurrent abscesses and boils in the armpits and groin. As the disease progresses it has a disabling affect on patients lives, causing severe pain and scarring, interfering with day to day activities such as sitting, walking and arm movements. We know that a number of patients with HS have family members who also suffer with HS or related conditions (e.g. acne). In these patients the condition is likely to be inherited and caused by a single gene. Some genes for HS have now been identified, however there would appear to be more to find, and it is important to understand how changes in these genes result in HS.

The purpose of this study is to try and identify further genes that cause HS in patients and their families and try to understand how changes in known and any newly identified genes result in the disease . We hope that this work will help us to understand the cause of HS and help scientists develop new treatments for HS and related conditions.

#### **Why have I been chosen?**

You have been diagnosed with HS. Comparing genetic material from people who are affected with HS to unaffected individuals and the genetic material of their relatives (who may not necessarily be affected) may help researchers.

#### **Do I have to take part?**

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without explaining why. A decision to withdraw at any time or a decision not to take part, will not affect the standard of care you receive.

**What do I have to do?**

In order to participate in the study we will ask you to sign a consent form. You will be asked to attend the hospital for one appointment, which should take approximately 10 minutes. You may wish to be seen now, if you have the time, or make an appointment for a later date.

**What will happen to me if I take part?**

On attending your appointment, a study doctor will discuss the purpose of the study and will go through the consent procedure with you. You will be able to ask any questions about your participation.

- You will be asked to complete a short questionnaire about your family in particular details of any relatives you know who also suffer with HS or related skin conditions (e.g. acne). This should take approximately 10 minutes.
- You will then be asked to donate a blood sample or saliva sample for analysis of your genetic material (DNA). A blood sample will involve taking a small amount of blood (approximately 17mls) from your arm. A saliva sample involves spitting 2.5mls (approximately 1 teaspoon) into a plastic pot.
- If we do identify a possible causative gene in you then you may, at a later stage, be asked to attend one further appointment to obtain a hair sample (6 hairs only) and/ or a small skin biopsy. Alternatively, if at any stage you are scheduled to have surgical treatment we may be able to use some excised skin.

**What are the possible disadvantages of joining this study?**

There are no perceived risks of giving a sputum sample. Blood tests can sometimes be a little uncomfortable and can cause minor bruising or irritation at the site. In rare cases there may be local infection. There are no perceived risks of giving hair plucks. Small skin biopsies may result in bleeding, a 4mm scar and there is a very small risk of infection. There are no other perceived risks.

Our research team is comprised of doctors who are highly trained to carry out these procedures. You must contact us if you experience any problems.

**What are the benefits of taking part?**

The information obtained from this study will help future treatment of patients with HS and related conditions, but is unlikely to benefit you personally. You may help future generations by advancing what is known about HS, in particular what causes the disease, which will help scientists develop new treatments.

**What happens if there is a problem?**

If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed by negligence, then you may have grounds for legal action but you may have to pay your legal costs. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal

National Health Service complaints mechanism will be available to you. In the event of this, please contact the Patient Advisory Liaison Service (PALS) via telephone 02071888803 or visit in person by asking at the hospital reception. You can also contact Dr Andrew Pink on 020 71886277 and he will try his best to resolve the matter for you. His address is: St John's Institute of Dermatology, St. Thomas' Hospital, London, SE1 7EH.

**Will I receive any money for taking part in this study?**

You will not receive any financial benefits from this study.

•

**Will my taking part in the study be kept confidential?**

All information that is collected about you during the course of the research will be kept strictly confidential. Only the research staff and approved regulatory personnel will have access to any information about you. Any information about you that is seen by other people will have your name and address removed so that you cannot be recognised.

**What will happen to the results at the end of the study**

Your samples will be stored for an indefinite period of time and will be potentially used for further research in HS and related diseases (acne). The samples will be stored anonymously. If you do not wish us at any time to continue to store your sample, then please contact the unit and we will provide you with a form to complete and return to us asking us to destroy your sample.

When the study has been completed we will aim to publish the results in a medical journal. Please let us know if you wish to be informed of the publication of the study and we will aim to keep you fully informed. You will not be identified in any publication.

**Who is organizing and funding the research?**

This study is being funded, organised and conducted by the Skin Therapy Research Unit (King's College London), Guy's and St Thomas' Hospitals London and by the Medical Research Council. As we recruit more participants and the project expands, we will apply to external funding sources.

**Who has reviewed the study?**

This study has been reviewed by independent clinicians and scientists and has been approved by the London Stanmore Research Ethics Committee.

**Who can I contact for further information on the study?**

Please contact Dr. Andrew Pink on 02071886277 any time if you have questions about the study.

**What will happen if I wish to withdraw from the study?**

Your participation is voluntary. You may refuse to be in this study, or withdraw from the study at any time without giving a reason. We would like to keep any samples you have already given us and continue to use them in our study.

None of these actions will affect your future treatment. If you decide to participate, your doctor can decide to withdraw you from the study at any time.

**If you have understood all the information above and wish to participate in the study you will be asked to sign an Informed Consent Form. You should keep a copy of this Information Sheet for yourself.**

Thankyou for your help and support

## **Information Leaflet: Family member**

### **Genetics Studies of Hidradenitis Suppurativa**

#### **Our invitation**

You are being invited to take part in a research study. Before you decide it is important to understand why the research is being done and what it will involve. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

#### **What is the purpose of the study?**

Hidradenitis Suppurativa (HS) is a chronic skin disease in which patients suffer with recurrent abscesses and boils in the armpits and groin. As the disease progresses it has a disabling affect on patients lives, causing severe pain and scarring, interfering with day to day activities such as sitting, walking and arm movements. We know that a number of patients with HS have family members who also suffer with HS or related conditions (e.g. acne). In these patients the condition is likely to be inherited and caused by a single gene. Some genes for HS have now been identified, however there would appear to be more to find, and it is important to understand how changes in these genes result in HS.

The purpose of this study is to try and identify further genes that cause HS in patients and their families and try to understand how changes in known and any newly identified genes result in the disease . We hope that this work will help us to understand the cause of HS and help scientists develop new treatments for HS and related conditions.

#### **Why have I been chosen?**

Your relative has been diagnosed with HS. Comparing genetic material from people who are affected with HS to that of their relatives (who may not necessarily be affected) will help researchers.

#### **Do I have to take part?**

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without explaining why. A decision to withdraw at any time or a decision not to take part, will not affect the standard of care you receive.

#### **What do I have to do?**

In order to participate in the study we will ask you to sign a consent form. You will be asked to attend the hospital for one appointment (approximately 15-20 minutes). In you are unable to attend our hospital we will try and arrange a telephone discussion instead (approximately 10 minutes).

### **What will happen to me if I take part?**

On attending your appointment, a study doctor will discuss the purpose of the study and will go through the consent procedure with you. You will be able to ask any questions about your participation.

- You will be asked to complete a short questionnaire about your medical history, your family's medical history, in particular whether you suffer with symptoms of HS or related skin conditions (e.g. acne). Your skin will be briefly checked to confirm if you have HS or a related condition. This should take approximately 10-15 minutes.
- You will then be asked to donate a blood sample or saliva sample for analysis of your genetic material (DNA). A blood sample will involve taking a small amount of blood (approximately 17mls) from your arm. A saliva sample involves spitting 2.5mls (approximately 1 teaspoon) into a plastic pot.
- If we do identify a possible causative gene in you then you may, at a later stage, be asked to attend one further appointment to obtain a hair sample (6 hairs only) and/ or a small skin biopsy.

*If you are unable to attend Guy's hospital for the questionnaire and skin check, we will arrange a telephone interview (approximately 10 minutes). If you unable to attend Guy's hospital for a blood/ saliva sample, we will post a simple saliva collection pot with instructions to your home.*

### **What are the possible disadvantages of joining this study?**

There are no perceived risks of giving a sputum sample. Blood tests can sometimes be a little uncomfortable and can cause minor bruising or irritation at the site. In rare cases there may be local infection. There are no perceived risks of giving hair plucks. Small skin biopsies may result in bleeding, a 4mm scar and there is a very small risk of infection. There are no other perceived risks.

Our research team is comprised of doctors who are highly trained to carry out these procedures. You must contact us if you experience any problems.

### **What are the benefits of taking part?**

The information obtained from this study will help future treatment of patients with HS and related conditions, but is unlikely to benefit you personally. You may help future generations by advancing what is known about HS, in particular what causes the disease, which will help scientists develop new treatments.

### **What happens if there is a problem?**

If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed by negligence, then you may have



grounds for legal action but you may have to pay your legal costs. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanism will be available to you. In the event of this, please contact the Patient Advisory Liaison Service (PALS) via telephone 02071888803 or visit in person by asking at the hospital reception. You can also contact Dr Andrew Pink on 020 71886277 and he will try his best to resolve the matter for you. His address is: St John's Institute of Dermatology, St. Thomas' Hospital, London, SE1 7EH.

**Will I receive any money for taking part in this study?**

You will not receive any financial benefits from this study.

•

**Will my taking part in the study be kept confidential?**

All information that is collected about you during the course of the research will be kept strictly confidential. Only the research staff and approved regulatory personnel will have access to any information about you. Any information about you that is seen by other people will have your name and address removed so that you cannot be recognised.

**What will happen to the results at the end of the study**

Your samples will be stored for an indefinite period of time and will be potentially used for further research in HS and related diseases (acne). The samples will be stored anonymously. If you do not wish us at any time to continue to store your sample, then please contact the unit and we will provide you with a form to complete and return to us asking us to destroy your sample.

When the study has been completed we will aim to publish the results in a medical journal. Please let us know if you wish to be informed of the publication of the study and we will aim to keep you fully informed. You will not be identified in any publication.

**Who is organizing and funding the research?**

This study is being funded, organised and conducted by the Skin Therapy Research Unit (King's College London), Guy's and St Thomas' Hospitals London and by the Medical Research Council. As we recruit more participants and the project expands, we will apply to external funding sources.

**Who has reviewed the study?**

This study has been reviewed by independent clinicians and scientists and has been approved by the London Stanmore Research Ethics Committee.

**Who can I contact for further information on the study?**

Please contact Dr. Andrew Pink on 02071886277 any time if you have questions about the study.

**What will happen if I wish to withdraw from the study?**

Your participation is voluntary. You may refuse to be in this study, or withdraw from the study at any time without giving a reason. We would like to keep any samples you have already given us and continue to use them in our study.

None of these actions will affect your future treatment. If you decide to participate, your doctor can decide to withdraw you from the study at any time.

**If you have understood all the information above and wish to participate in the study you will be asked to sign an Informed Consent Form. You should keep a copy of this Information Sheet for yourself.**

Thankyou for your help and support

## Genetic Studies of Hidradenitis Suppurativa **CONSENT FORM: Patient**

LREC Study Number:

\_\_\_\_\_  
Name of participant

\_\_\_\_\_  
Identification Number

Name of Investigators: Professor J Barker, Dr A Pink

**Please initial box**

1. I confirm that I have read and understand the information sheet (Version dated 3/6/2011) for the above study for hidradenitis patients and their family members and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of any of my medical notes may be looked at by hospital staff and regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐
4. I agree that my family members can be contacted to participate in this study once I have discussed it with them. ☐
5. I agree to take part in the above study. ☐

\_\_\_\_\_  
Signature\*

\_\_\_\_\_  
Name of participant\*

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name of researcher

\_\_\_\_\_  
Date

\*Relationship to participant (if participant unable to sign) \_\_\_\_\_

### **Statement of Interpreter** (where appropriate)

I have interpreted the information above to the best of my ability and in a way which I believe he/she can understand.

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name (PRINT)

\_\_\_\_\_  
Date

## Genetic Studies of Hidradenitis Suppurativa **CONSENT FORM: Family member**

LREC Study Number:

\_\_\_\_\_  
Name of participant

\_\_\_\_\_  
Identification Number

Name of Investigators: Professor J Barker, Dr Andrew Pink

**Please initial box**

1. I confirm that I have read and understand the information sheet (Version dated 3/6/2011) for the above study for hidradenitis patients and their family members and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of any of my medical notes may be looked at by hospital staff and regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐
4. I agree to take part in the above study. ☐

\_\_\_\_\_  
Signature\*

\_\_\_\_\_  
Name of participant\*

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name of researcher

\_\_\_\_\_  
Date

\*Relationship to participant (if participant unable to sign)  
\_\_\_\_\_

**Statement of Interpreter** (where appropriate)

I have interpreted the information above to the best of my ability and in a way which I believe he/she can understand.

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name (PRINT)

\_\_\_\_\_  
Date

**CONSENT TO DONATION AND STORAGE OF BLOOD OR SALIVA  
SAMPLES FOR FUTURE MEDICAL RESEARCH IN HIDRADENITIS  
SUPPURATIVA AND RELATED DISEASES**

Name of investigators and department:	Professor J Barker, Dr Andrew Pink, Guys and St. Thomas' Hospital London
Research Ethics Committee Reference:	
Sample to be taken:	BLOOD and/or SALIVA +/- HAIR PLUCKS and/or SKIN BIOPSY

Blood, saliva, hair or skin collected as part of the research project may be stored and used by St. John's Institute of Dermatology (Guys & St. Thomas' Hospital London) for future medical research. All personal data collected for the research project will be anonymised. You will receive the outcome of future research projects if requested. All staff undertaking future studies will abide by the Data Protection Act 1998 with any medical information relating to you being kept confidential. All samples will be disposed of lawfully when no longer required.

I understand that the blood/ saliva/ hair or skin samples taken for this study will not be used for diagnostic purposes. ☐

I accept that I have given my consent voluntarily for the storage of blood/ saliva/ hair or skin samples and that I am free to withdraw my consent at any time. ☐

I agree that the blood/ saliva/ hair/ skin may be used for future genetic research but not for research that involves reproductive cloning without my express consent. ☐

I agree that following completion of this study my personal data and materials may be stored for future studies in hidradenitis suppurativa in a Research Tissue Bank. This will only take place following the necessary ethical approval required to establish a Research Tissue Bank. ☐

If during the course of the study, I lose the capacity to consent, any data or samples previously obtained with my consent will be retained and used in the study. ☐

If you have any comments please include them here:

\_\_\_\_\_  
Signature of participant

\_\_\_\_\_  
Name (PRINT)

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature of investigator

\_\_\_\_\_  
Name (PRINT)

\_\_\_\_\_  
Date

# Patient Clinical Research Form

## PATIENT DETAILS

### PATIENT DEMOGRAPHICS

Name.....

DoB .....

Male/ female .....

Occupation.....

Recruitment Source:

St George's clinic

☐

St John's clinic

Other

.....

### ANTHROPOMETRY

Weight .  Kg Height .  cm BMI

Waist circumference .  cm Blood pressure  mm/Hg

### RACE

White

☐

White Irish

☐

Asian

☐

Black

☐

Chinese

☐

Mixed race, please state .....

☐

Other, please state .....

☐

### SMOKING

Current smoker

Yes

☐

No

☐

Ex-smoker (smoked at any stage in the past)

Yes

☐

No

☐

Start smoking prior to disease development?

Yes

☐

No

☐

(If so, how long before?  Years )

Pack/year history (1 pack year history = 20 cigarettes smoked for 1 year)

## CLINICAL HISTORY

### HS

History of 5 or more boils/abscesses in flexural sites

YES ☐ NO ☐

Which sites have been affected (include present or past):

Posterior ears ☐  
 Neck ☐  
**Axilla** ☐  
 Arms ☐  
**Submammary** ☐  
 Chest wall ☐  
 Back ☐  
 Suprapubic ☐  
**Groin** ☐  
 Genital ☐  
**Buttocks** ☐  
 Thighs ☐  
 Legs ☐  
 Other ..... ☐

Age when first affected

Frequency of flares

$\geq 12$ /year ☐ 6-11/year ☐ 1-5/year ☐  $\leq 1$ /year ☐

Related to menstrual cycle

YES ☐ NO ☐

(If so, how long before menstruation on average?)

days

Systemic upset associated with flares  
(fever/malaise/fatigue)

YES ☐ NO ☐

Requiring hospital admission at any stage

YES ☐ NO ☐

## PAST MEDICAL HISTORY

Acne vulgaris	<input type="checkbox"/>	
Nodulocystic acne	<input type="checkbox"/>	
Scalp folliculitis	<input type="checkbox"/>	
Dissecting cellulitis	<input type="checkbox"/>	
Pilonidal sinus	<input type="checkbox"/>	
Squamous cell carcinoma	<input type="checkbox"/>	
Psoriasis	<input type="checkbox"/>	
Eczema	<input type="checkbox"/>	
Keratosis pilaris	<input type="checkbox"/>	
Acanthosis nigricans	<input type="checkbox"/>	
History of any other skin disorder   YES <input type="checkbox"/> NO <input type="checkbox"/>		
If so, please detail .....		
Keratitis	<input type="checkbox"/>	
Crohn's disease	<input type="checkbox"/>	
Ulcerative colitis	<input type="checkbox"/>	
Arthritis	<input type="checkbox"/>	Diagnosis .....
Diabetes	<input type="checkbox"/>	(Type 1 <input type="checkbox"/> Type 2 <input type="checkbox"/> )
PCOS	<input type="checkbox"/>	
Alzheimers disease	<input type="checkbox"/>	



**OTHER PAST MEDICAL HISTORY**

Please fill in only one diagnosis per line.

SD No.	Diagnosis	Onset Date		Date Resolved	Tick if Ongoing
		M	Yr		
1		<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
2		<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
3		<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
4		<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
5		<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		

**DRUG HISTORY****CURRENT THERAPY FOR HIDRADENITIS SUPPURATIVA**

Treatment Name	Start Date		Responder Status (Responder, indeterminate response, non-responder*)	Ongoing (tick)
	Month	Year		
1.	<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
2.	<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
3.	<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
4.	<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		

\* Responder = Patient felt there was a **definite** clinical improvement on this medication

Indeterminate = Patient unsure whether the medicine resulted in any clinical improvement

Non-responder = Patient reports that the medication had **no** effect

## OTHER CONCOMITANT MEDICATION (on today's visit)

Treatment Name	Start Date	Indication (Should relate to a medical diagnosis stated above)	Ongoing (tick)
	Month    Year		
1.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
2.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
3.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
4.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
5.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
6.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
7.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
8.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
9.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
10.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		
11.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>		

## PREVIOUS THERAPY FOR HIDRADENITIS SUPPURATIVA

(topical/ oral antibiotics/ acitretin/ isotretinoin / ciclosporin/ dapsone/ anti-TNF / **surgery** / other)

Treatment Name	Start Date Stop Date	Responder Status (responder, indeterminate or non-responder)
	Month    Year	
1.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	
2.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	
3.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	
4.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	

## **FAMILY SKIN DISEASE HISTORY**

This would include hidradenitis suppurativa or other related diseases i.e. acne vulgaris, conglobate acne, dissecting folliculitis of the scalp, pilonidal sinus, folliculitis, atopic dermatitis, psoriasis or any other skin disease.

Anyone else in the family affected with HS

Yes ☐

No ☐

(Tick box if family history UNKNOWN ☐ )

If so, please draw family pedigree:

### **ANY FAMILY HISTORY OF**

Acne vulgaris

☐

Scalp folliculitis/ Dissecting cellulitis

☐

Pilonidal sinus

☐

Squamous cell carcinoma

☐

Alzheimers

☐

## EXAMINATION/ CLINICAL MEASURES

LESION TYPE	Comedones	Nodules	Sinus tracts	Scarring	Disease active	
					YES	NO
Axilla	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Submammary	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Suprapubic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Groin	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Genital	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Buttocks	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acanthosis nigricans	<input type="checkbox"/>	Where?.....				
Keratosis pilaris	<input type="checkbox"/>	Where?.....				
Lymphoedema	<input type="checkbox"/>	Where?.....				
Acne vulgaris	<input type="checkbox"/>					
Nodulocystic acne	<input type="checkbox"/>					
Acne conglobata	<input type="checkbox"/>					
Scalp folliculitis/ dissecting cellulitis		<input type="checkbox"/>				

**GRADING SCORES**

HURLEY STAGE (Please ring as appropriate)

I       $\geq$  1 inflammatory nodule with NO sinuses or scarring

II      $\geq$  1 widely separated inflammatory nodules + scarring + sinuses

III    Multiple interconnected nodules + sinuses throughout entire affected area

Sartorius score

**QUALITY OF LIFE**

Dermatology Life Quality Index (DLQI) Score  (patient should be given sheet to complete)

Time from first symptoms to presentation to a Dermatologist  years  months

From where were they referred/ where was it picked up:

GP ☐

A&E ☐

General surgery ☐

Other .....

**BLOOD TESTS**

Cholesterol

Triglycerides

HDL

Fasting glucose

CRP

ESR

**CULTURES**

Bacteria cultured

Type	Dates

**MRI performed**

YES

☐

NO

☐

If yes:

Site:.....

Result:.....

**BLOOD/SALIVA SAMPLING**

Date Collected

Blood

Yes ☐No ☐

Saliva

Yes ☐No ☐**DNA Extraction**Yes ☐No ☐**DNA Number:****THERAPIES TO BE COMMENCED TODAY**

Treatment Name	Start Date
	Month Year
1.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
2.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
3.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
4.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>

<b>OVERALL FEELING re: PHENOTPYE</b> .....  Form completed by .....  Job/Title .....
--

**Advertisement for recruitment of volunteers for the Molecular Genetic Analysis of Hidradenitis Suppurativa study. Ref: 11/LO/0966 approved by the NRES Committee London-Stanmore on 26<sup>th</sup> August 2011.**

You are under no obligation to take part in this study however if you choose to, participation in this research is voluntary and you may withdraw at any time.

We would like to invite you to participate in a study in which we are investigating the genetic basis of a skin condition called Hidradenitis Suppurativa. We hope that the information obtained from analysis of these samples will help us better understand the processes involved in this chronic inflammatory skin condition and in the future therapies.

You will be asked to provide 3 types of samples:

- 17 mls (approximately four level teaspoons) of your blood for DNA analysis and for the study your blood cells.
- Six hair plucks from your scalp for the study of your skin.
- Two 6mm punch biopsy samples, one from a flexural area (eg. arm pit) and one from a non-flexural area (eg. back).
- In addition you will be asked some questions about yourself – gender; age; ethnicity; medical history (regarding skin disease)

The 2 punch biopsy sites will each have 2 stitches in place for 10 days. You will need to return to the clinic to have these stitches removed.

You can take part if:

- You are over 18
- You have no medical history of hidradenitis suppurativa, pilonidal abscesses or acne

You will receive compensation for your time **of £150**. Any risks involved will be detailed in the information sheet. This is not a treatment study. All participant information will be anonymised and kept confidential.

If you are interested in participating please contact one of the study team for more detailed information: Andrew.pink@kcl.ac.uk Tel Nos: **0207 188 6412**.

## **Information Leaflet: Volunteer**

### **Genetics Studies of Hidradenitis Suppurativa**

#### **Our invitation**

You are being invited to take part in a research study. Before you decide it is important to understand why the research is being done and what it will involve. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

#### **What is the purpose of the study?**

Hidradenitis Suppurativa (HS) is a chronic skin disease in which patients suffer with recurrent abscesses and boils in the armpits and groin. As the disease progresses it has a disabling affect on patients lives, causing severe pain and scarring, interfering with day to day activities such as sitting, walking and arm movements. We know that a number of patients with HS have family members who also suffer with HS or related conditions (e.g. acne). In these patients the condition is likely to be inherited and caused by a single gene. Some genes for HS have now been identified, however there would appear to be more to find, and it is important to understand how changes in these genes result in HS.

The purpose of this study is to try and identify further genes that cause HS in patients and their families and try to understand how changes in known and any newly identified genes result in the disease . We hope that this work will help us to understand the cause of HS and help scientists develop new treatments for HS and related conditions.

#### **Why is it important for unaffected volunteers to take part in this study?**

Comparing genetic material, cells and skin from people who are affected with HS with that of unaffected healthy volunteers will help researchers to understand the key differences that result in disease.

#### **Do I have to take part?**

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without explaining why. A decision to withdraw at any time or a decision not to take part will not affect the standard of medical care that you receive now or in the future.

#### **What do I have to do?**



In order to participate in the study we will ask you to sign a consent form. You will be asked to attend the hospital for an initial appointment lasting approximately 60 minutes and then for a follow-up appointment 10 days later lasting approximately 10 minutes.

**What will happen to me if I take part?**

On attending your appointment, a study doctor will discuss the purpose of the study and will go through the consent procedure with you. You will be able to ask any questions about your participation.

- You will be asked to complete a short questionnaire about your medical history, in particular whether you suffer with symptoms of HS or related skin conditions (e.g. acne).
- You will then be asked to donate a blood sample. This will involve taking a small amount of blood (approximately 17mls) from your arm.
- You will then be asked to donate a hair sample (6 hairs, plucked from your scalp) and two small skin biopsies (one from your armpit, groin or buttock and one from nearby skin). The skin biopsies will be performed under local anaesthetic. You will be asked to return 10 days later for the removal of stitches.

**What are the possible disadvantages of joining this study?**

Blood tests can sometimes be a little uncomfortable and can cause minor bruising or irritation at the site. In rare cases there may be local infection. There are no perceived risks of giving hair plucks. Small skin biopsies may result in bleeding, a 6mm scar and there is a very small risk of infection. There are no other perceived risks. Our research team is comprised of doctors who are highly trained to carry out these procedures. You must contact us if you experience any problems.

**What are the benefits of taking part?**

The information obtained from this study will help future treatment of patients with HS and related conditions but is unlikely to benefit you personally. You may help future generations by advancing what is known about HS, in particular what causes the disease, which will help scientists develop new treatments.

**What happens if there is a problem?**

If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed by negligence, then you may have grounds for legal action but you may have to pay your legal costs. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanism will be available to you. In the event of this, please contact the Patient Advisory Liaison Service (PALS) via telephone 02071888803 or visit in person by asking at the hospital reception. You can also contact Dr Andrew Pink on 020 71886277 and he will try his best to

resolve the matter for you. His address is: Floor 9, Guy's Tower, Guy's Hospital, London, SE1 7EH.

**Will I receive any money for taking part in this study?**

You will receive £150 for taking part in this study.

•

**Will my taking part in the study be kept confidential?**

All information that is collected about you during the course of the research will be kept strictly confidential. Only the research staff and approved regulatory personnel will have access to any information about you. Any information about you that is seen by other people will have your name and address removed so that you cannot be recognised.

**What will happen to the results at the end of the study**

Your samples will be stored for an indefinite period of time and will potentially be used for further research in HS and related diseases (acne). The samples will be stored anonymously. If you do not wish us at any time to continue to store your sample, then please contact the unit and we will provide you with a form to complete and return to us asking us to destroy your sample.

When the study has been completed we will aim to publish the results in a medical journal. Please let us know if you wish to be informed of the publication of the study and we will aim to keep you fully informed. You will not be identified in any publication.

**Who is organizing and funding the research?**

This study is being funded, organised and conducted by the Skin Therapy Research Unit (King's College London), Guy's and St Thomas' Hospitals London and by the Medical Research Council.

**Who has reviewed the study?**

This study has been reviewed by independent clinicians and scientists and has been approved by the London Stanmore Research Ethics Committee.

**Who can I contact for further information on the study?**

Please contact Dr. Andrew Pink on 02071886277 any time if you have questions about the study.

**What will happen if I wish to withdraw from the study?**

Your participation is voluntary. You may refuse to be in this study, or withdraw from the study at any time without giving a reason. We would like to keep any samples you have already given us and continue to use them in our study.

None of these actions will affect your future treatment. If you decide to participate, your doctor can decide to withdraw you from the study at any time.

**If you have understood all the information above and wish to participate in the study you will be asked to sign an Informed Consent Form. You should keep a copy of this Information Sheet for yourself.**

Thankyou for your help and support

## Genetic Studies of Hidradenitis Suppurativa **CONSENT FORM: Volunteer**

LREC Study Number:

\_\_\_\_\_  
Name of participant

\_\_\_\_\_  
Identification Number

Name of Investigators: Professor J Barker, Dr Andrew Pink

**Please initial box**

1. I confirm that I have read and understand the volunteer information sheet (Version dated 16/5/2012) for the above study and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of any of my medical notes may be looked at by hospital staff and regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. ☐
4. I agree to take part in the above study. ☐

\_\_\_\_\_  
Signature\*

\_\_\_\_\_  
Name of participant\*

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name of researcher

\_\_\_\_\_  
Date

\*Relationship to participant (if participant unable to sign) \_\_\_\_\_

### **Statement of Interpreter** (where appropriate)

I have interpreted the information above to the best of my ability and in a way which I believe he/she can understand.

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name (PRINT)

\_\_\_\_\_  
Date

# Volunteer Clinical Research Form

Please tick YES or NO (if the answer is NO to any of these questions then the patient is to be excluded from the study)

INCLUSION CRITERIA	YES	NO
Written informed consent		
≥ 16 years		

Please tick YES or NO (if the answer is YES to any of these questions then the patient is to be excluded from the study)

EXCLUSION CRITERIA	YES	NO
Unable to give written informed consent		
< 16 years old		
Blood transfusion received within last 4 weeks		
History of HS		
Family history of HS		
History of related disorder of follicular occlusion (acne, pilonidal abscess, dissecting cellulitis)		

**Consent obtained** (forms 1 and 2)

Yes ☐ No ☐

•

• Name..... DoB ..... Male/ female

.....

Weight 

--	--	--

 . 

--

 Kg

Height 

--	--	--

 . 

--

 Cm

BMI

--	--

Current smoker

Yes ☐ No ☐

Ex-smoker (smoked at any stage in the past)

Yes ☐ No ☐

Pack/year history (1 pack year history = 20 cigarettes smoked for 1 year)

--	--

White

--

White Irish

--

Asian

--

Black

--

Chinese

--

Mixed race, please state .....

Other, please state .....

--

Previous history of: Acne vulgaris

Scalp folliculitis/ dissecting cellulitis

Pilonidal sinus

Crohn's disease


Is any one else in the family affected

Yes ☐ No ☐

If so, please draw family pedigree

## Sample collection

Date Collected 

--	--

 / 

--	--

 / 

--	--	--	--

Form completed by ..... Job/Title .....

### Appendix 3 – Primers used to amplify exons of *NCSTN*, *PSENE1* and *PSEN1*

Primers used to amplify all exons of *NCSTN*, *PSENE1* and *PSEN1*:

Gene	Exon	Forward primer	Reverse primer
<i>NCSTN</i>	1	CGAACTTCCGGTCTCTTAGG	CAGGAACCAGAAGGGAGACA
<i>NCSTN</i>	2	CCCAAGCCATACTTCAGGTT	TCTCCTAAGCAGGACATGGA
<i>NCSTN</i>	3	GGGCAACAGCTTTTCAGTTC	CAGCCAAGGGAAGTTTTAC
<i>NCSTN</i>	4	CCCCTAGTTCCCCATTTGTT	CATCCCTCACAAAGACGTAAG
<i>NCSTN</i>	5	AAGATTCTTCTGCCGTCACC	TGCCATCTCAGCTCACTGAA
<i>NCSTN</i>	6	ATGTCTCCGAAAAGGGTGTG	ATGCAGAGCTCCTTCATGGT
<i>NCSTN</i>	7	GGCACTGGTCAGAGATTTCC	GGCTTCCCAATTTAGCAACT
<i>NCSTN</i>	8	CCTCAAATGGGGAGGAATCT	GCTGAGGTCTTCTCCAACCA
<i>NCSTN</i>	9	CCTACAGCTTTGATGATCTG	GAAGAGCGGCTGCTTCCT
<i>NCSTN</i>	10	CGCCTCTTCCCTTGACTTCT	CTGGGGCTTGGTGGAAC
<i>NCSTN</i>	11	ATAGGGTAGCTCCCCAAGCA	GTTCTGGAGGGGGTAGTGTG
<i>NCSTN</i>	12	TCCCTGCATAAGGAGCATTT	AATGGCCACCTCTCCTATC
<i>NCSTN</i>	13	CACTGCCAAGGTAGCACTGA	GTTTGGGACACTAGGGCAAG
<i>NCSTN</i>	14	TATCTGGCCAGTCTGGTTCC	GAAGGGCAATGGAAAAGGAT
<i>NCSTN</i>	15	TTCCATTGCCCTTCTTTCTG	AACACCCAATTGCTGAGGAC
<i>NCSTN</i>	16	TTGGTTAGAGCATGGCACTG	ATGGGAGATGGAAAGGGCTA
<i>NCSTN</i>	17	GGAGAGATGTTGCCCATGAT	AGGGAGGTTCCAGTGACAAA
<i>PSEN1</i>	1	TGCACAAAGTTCTGTTTTCTTC	CTAGCCCCCTCAGCATTTCT
<i>PSEN1</i>	2	TGACGGGTCTGTTGTAAATCC	CTGCTCCTGACCATCACAGA
<i>PSEN1</i>	3	AAAATTCTGTGTTGGAGGTGGT	TTACACATGCACCTGGCTTC
<i>PSEN1</i>	4	GCGACAAAGTGAGACCCTGT	CAACAAAGTACATGGCTTTAAATGA
<i>PSEN1</i>	5	TGTTTGGGAGCCATCACAT	CCAGCCGAAATCTTCAAATG
<i>PSEN1</i>	6	TCGTTAATTCTCCCTACCA	CAGGAATGCTGTGCATTTAAGA
<i>PSEN1</i>	7	GGAAGACTGGCGATTTGTGT	GGAGTCTATGACCAAAGAAAGACG
<i>PSEN1</i>	8	ACTGCTACAGCCCATGCTTT	CACTCTCAATTCATTTTATTCTCAA
<i>PSEN1</i>	9	GGGTTGAGTAGGGCAGTGATA	CTGGGGTTAAACAGAACTGC
<i>PSEN1</i>	10	TTCCAGATTGAATGAACGTCTG	TGGACACAGGAAAATCACCTT
<i>PSENE1</i>	1	TTCTAATAGGGGCGTGTTG	TCGCAGGTCCTTCATCTCTT
<i>PSENE1</i>	2	ATCCCAAAGAGGAGCCAGAT	CCTAGACCAGCCTTCCCTTC
<i>PSENE1</i>	3	GAGAGGGGGAAGCGGGGAAT	GCCTCGTACAAAGCTCTCAACA

## Appendix 4 – Left and right hybridisation sequences used for MLPA

### Left hybridisation sequences for MLPA

<i>PSENE1</i> 1	GGGTTCCCTAAGGGTTGGAGACGACCCGGCCCCAGCGCAGCTATGAACCTGGAGCGAGTGT
<i>PSENE1</i> 2	GGGTTCCCTAAGGGTTGGATTCTGCCTTTTCTCTGGTTGGTCAACATCTTCT
<i>PSENE1</i> 3	GGGTTCCCTAAGGGTTGGATGTCCTGGGAGCCCCCTTCTCAAACCTCCTAAGACTTGT
<i>NCSTN</i> 1	GGGTTCCCTAAGGGTTGGAGAAACACGAACCTCCGGTCTCTTAGGCTCCGGGCCACAGAGACGGTGTCAAGTGGTA
<i>NCSTN</i> 2	GGGTTCCCTAAGGGTTGGATGCAGGGGAAACTCAGTGGAGAGGAAGATATATATCCCCCTTA
<i>NCSTN</i> 3	GGGTTCCCTAAGGGTTGGATTGACTGATGGCCCCAACCCCCCTTA
<i>NCSTN</i> 4	GGGTTCCCTAAGGGTTGGATTAATGGAGAAAGCTGAAAGGGAGAACCGCCGAAT
<i>NCSTN</i> 5	GGGTTCCCTAAGGGTTGGATTCTATGGGCCAGAGTTTGCTCACTGCAGAGAAATACAGTGGAAAT
<i>NCSTN</i> 6	GGGTTCCCTAAGGGTTGGATGCCATGCAGCTCTTTTCACACATGCATGCTGTCTAT
<i>NCSTN</i> 7	GGGTTCCCTAAGGGTTGGATGTCTGATTACAATGTGTGGAGCATGCTAAAAG
<i>NCSTN</i> 8	GGGTTCCCTAAGGGTTGGATTTTCTGGAATGTGGCCCCAGGGGCTGAAAGCGCAGTGGCTTCCTTT
<i>NCSTN</i> 9	GGGTTCCCTAAGGGTTGGACATTGGCAGCTCGAGGATGGTCTA
<i>NCSTN</i> 10	GGGTTCCCTAAGGGTTGGAGAGCTTTGGATGCACACAGATCCTGTTT
<i>NCSTN</i> 11	GGGTTCCCTAAGGGTTGGAGAGTGGTGTCTGGTGTCCCTGCTGTCATCCTCAGGAGGCCAAATC
<i>NCSTN</i> 12	GGGTTCCCTAAGGGTTGGATTACCAGAGTATTTACGACACTGCTGAGAACATTAATGTG
<i>NCSTN</i> 13	GGGTTCCCTAAGGGTTGGAGATGTGGCCACGGTGTCTGGGACGTGCTCTGTATG
<i>NCSTN</i> 14	GGGTTCCCTAAGGGTTGGATATGGGTTCTGATTAAAGCCAACAACCTCA
<i>NCSTN</i> 15	GGGTTCCCTAAGGGTTGGATCAACATTACATCGCTGTCTCCAGCCCCACCAACACCACTTATGTTGTAC
<i>NCSTN</i> 16	GGGTTCCCTAAGGGTTGGAGTCCAGGGCCCTTTGCATTCTAATGAGACGGACCGACTCCCCCGGTGTGT
<i>NCSTN</i> 17	GGGTTCCCTAAGGGTTGGATTCACTTCCTAGAGCATCTGTCCCACTGGGACACAACCACTAATTTGTCACTGG

### Right hybridisation sequences for MLPA

<i>PSENE1</i> 1	CCAATGAGGAGAAATTGAACCTGTGCCGGAAGTACTACTCTAGATTGGATCTTGCTGGCAC
<i>PSENE1</i> 2	GGTCTTCCGAGAGGCCCTTCCTTGTCAGCCTATCTAGATTGGATCTTGCTGGCAC
<i>PSENE1</i> 3	TTTCATGTCCACGTTCTCTGCTGACATCCCCCAATATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 1	GCCTAGAGAGGCCGCTAACAGACAGGAGCCGAACGGGGGCTTCCGCTCAGCAGAGATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 2	AATAAAACAGCTCCCTGTGTTGCGCTGCTCAACGCCACTCATTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 3	CATGGTTCTGCTGGAGAGCAAGCATTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 4	TGCTGGTCTTGCACTGTCCTTGACCAAGCCAGTCCTGCCTTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 5	TCGCTGGGCAATGGTTTGGCTTATGAAGACTTTAGTTTCCCCATCTTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 6	CAGCACTGCCACCTGCATGCGGCGCAGCTCCATCCATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 7	CCTATAAATACAACCTGGGACATTAAGCCTGATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 8	GTCACCCAGCTGGCTGCTGCTGAAGCTTTGCAAAAGGCACCTGATGTGTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 9	CGATATGGAGAAGGGCAAGTTTCCTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 10	CTCAGAAAAATGAGTCTGTACGGAACCATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 11	AGTCCCAGCCTCTCCACCATCTTCCCTGCAGCGATTCTTCGATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 12	AGCTATCCCGAATGGCTGAGCCCTGAAGAGGACCTGAACCTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 13	AGCTTGCAGGAGGAACCAACTTCAGCGACACAGTTCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 14	TGGTTCCAGTCTATCCTCAGGCAGGACCTATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN</i> 15	AGTATGCCTTGGCAAATTTGACTGGCACAGTGGTCAACCTCACCCGAGAGTCTAGATTGGATCTTGCTGGCAC

<i>NCSTN 16</i>	GCGTTCTACTGCACGATTAGCCAGGGCCTTGTCCTGCCTTTGAACTGATCTAGATTGGATCTTGCTGGCAC
<i>NCSTN 17</i>	AACCTCCCTGGGCCTGTCTCAGATTGGGATTAACATAAAAGAGTGGAACATCCTCTAGATTGGATCTTGCTGGCAC



## Appendix 5 – Regions of linkage in pedigrees 1 – 6 (reported in Chapter 6)

### Family 1

Chromosome	From (bp)	To (bp)
1	229355000	232007196
3	36214197	45296740
3	51568137	58036651
3	59475736	68812466
3	70274229	74112675
3	76992952	85683470
3	107805334	115340900
4	35653192	37996334
6	122945140	127208390
6	129980436	130936249
6	134248060	137941227
6	159944996	166570346
7	44808091	49435299
7	52544356	53662097
9	1	1519043
9	2951475	6722212
10	130665155	132992902
11	18001888	20301930
12	5603632	6905181
13	79801863	105328137
16	605481	2982708
16	3747204	7442293
18	66881895	70501767
19	56372705	57182204
20	60091406	60803463

### Family 2

Chromosome	From (bp)	To (bp)
1	40049184	48048060
1	197482265	227418130
2	125061838	154452223
3	7142227	36979042
4	145658429	147246721
4	176721793	188905061
5	14989346	26809915
5	73986648	87513775
7	146226582	153970767
8	27405809	105805344
8	138467144	End

9	1	2196734
10	28294586	89912106
10	132992902	End
11	1	4639292
11	103506972	118755738
12	131225471	End
14	100087364	104155317
18	33674808	72022671
19	14450812	17927041
19	56822413	End
21	27514740	47417303

### **Family 3**

Chromosome	From (bp)	To (bp)
1	1	5347155
1	100376448	200107578
1	234524045	End
2	1	4050728
2	42012857	224713026
4	140888081	End
6	1	End
7	1	152402621
8	11728994	73379450
8	122805639	132007359
10	51503335	End
11	16007446	75390240
12	20605116	79391433
13	31587270	67798928
15	76402624	End
16	Start	8088550
19	3030118	End
21	40708702	42896828

### **Family 4**

Chromosome	From (bp)	To (bp)
1	57760458	95276691
1	115427218	222093006
2	173686798	206830563
3	44141157	73336315
3	97741821	150609866
4	66773365	121429891
5	141573265	172737199
6	1	13097658
6	14492000	42658752
7	109297612	135354582

8	1	18207413
9	9501741	37020622
9	109312922	1315276676
10	1	7722848
10	125665080	End
11	5491183	12083579
11	107874113	113477081
12	78636050	117082846
13	106632499	End
14	94954368	End
15	58177757	92604461
16	80071458	End
17	6285864	12171195
17	35010976	End
18	23413944	46976586
20	43970377	50466232
22	1	43432826

### **Family 5**

Chromosome	From (bp)	To (bp)
2	15807637	27384579
2	181314519	192458747
5	174897200	End
6	1	1295151
6	40921160	43396325
6	109049607	136227558
6	148995849	163697439
7	33926351	37951726
7	38741626	45522229
9	1519043	4895391
9	78724005	81125839
10	49201964	60548276
10	82454295	132139508
11	18001888	66079786
11	74401763	85766185
12	10848803	13870080
12	122318985	126171662
13	1	30670985
14	99157430	105154105
15	86278479	95378243
15	96856201	End
16	1	3747204
16	83727833	89587871
17	75038439	End
18	8439366	10597874
19	52712828	End

**Family 6**

Chromosome	From (bp)	To (bp)
1	1	106774642
1	154688197	End
2	66665146	80698371
2	141096291	165647332
3	1	3767186
3	30690424	64171993
3	116269592	169801658
3	183995341	End
4	1	56977309
4	71439354	112591777
4	152900692	End
5	1	14989346
5	30881510	71674855
5	74302658	116319808
5	162626996	End
6	137076010	End
7	139185449	End
8	19076995	48089751
9	1	74186310
9	111945992	End
10	15613080	90759724
10	132992902	End
11	1	3743046
11	89892250	128379964
12	1	12578786
12	89653020	97416064
12	127284785	End
13	37877935	74474851
14	1	26840301
15	1	60145962
15	96856201	End
16	1	8305679
16	13489119	75483806
17	1	3569913
18	60568428	End
19	36904463	End
20	2488753	End
21	1	27168049
21	42293410	End
22	45251423	48526700